

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:54:12 ; Search time 2718 Seconds

(without alignments)
14904.748 Million cell updates/sec

Title: US-09-816-467-3

Perfect score: 1392

Sequence: 1 atggttttttcaacacccaat.....aggagatgacaaatgatataa 1392

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

GenBank:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1387.4	99.7	1791	1	CLOTEFTOX	M12739 C.tetani te
2	1387.4	99.7	1851	6	E01135	E01135 CDNA encodi
3	1387.4	99.7	4338	1	CTTEROXG	X06214 Clostridium
4	1387.4	99.7	4400	1	CTTERANG	X04436 Clostridium
5	1379.8	99.1	1858	6	AR017561	AR017561 Sequence
6	1350.4	97.0	3931	1	AF389424	AF389424 Clostridi
7	1346.4	96.7	1359	6	A17527	A17527 human fragm
8	1346.4	96.7	1359	6	I28430	I28430 Sequence 1
9	1336.8	96.0	1356	1	AF154828	AF154828 Clostridi
10	1287.2	92.5	1359	6	I13981	I13981 Sequence 7
11	799.2	57.4	1359	6	I28431	I28431 Sequence 3
12	799.2	57.4	1359	6	I28431	I28431 Sequence 3
13	799.2	57.4	1359	6	I28431	I28431 Sequence 3
14	797.2	57.3	4366	6	A49987	A49987 Sequence 4
15	796.8	57.2	3754	6	A42484	A42484 Sequence 12
16	796.8	57.2	3754	6	A42478	A42478 Sequence 17
17	796.8	57.2	3754	6	A42478	A42478 Sequence 6
18	796.8	57.2	3754	6	A42478	A42478 Sequence 18
19	796.8	57.2	3754	6	A42478	A42478 Sequence 5
20	796.8	57.2	3754	6	A42478	A42478 Sequence 9
21	796.8	57.2	3754	6	A42478	A42478 Sequence 6
22	796.8	57.2	3754	6	A42478	A42478 Sequence 6
23	796.8	57.2	3754	6	A42478	A42478 Sequence 6
24	796.8	57.2	3754	6	A42478	A42478 Sequence 6
25	796.8	57.2	3754	6	A42478	A42478 Sequence 6
26	796.8	57.2	3754	6	A42478	A42478 Sequence 6
27	796.8	57.2	3754	6	A42478	A42478 Sequence 6
28	796.8	57.2	3754	6	A42478	A42478 Sequence 6
29	796.8	57.2	3754	6	A42478	A42478 Sequence 6
30	796.8	57.2	3754	6	A42478	A42478 Sequence 6
31	796.8	57.2	3754	6	A42478	A42478 Sequence 6
32	796.8	57.2	3754	6	A42478	A42478 Sequence 6
33	796.8	57.2	3754	6	A42478	A42478 Sequence 6
34	796.8	57.2	3754	6	A42478	A42478 Sequence 6
35	796.8	57.2	3754	6	A42478	A42478 Sequence 6
36	796.8	57.2	3754	6	A42478	A42478 Sequence 6
37	796.8	57.2	3754	6	A42478	A42478 Sequence 6
38	796.8	57.2	3754	6	A42478	A42478 Sequence 6
39	796.8	57.2	3754	6	A42478	A42478 Sequence 6
40	796.8	57.2	3754	6	A42478	A42478 Sequence 6
41	796.8	57.2	3754	6	A42478	A42478 Sequence 6
42	796.8	57.2	3754	6	A42478	A42478 Sequence 6
43	796.8	57.2	3754	6	A42478	A42478 Sequence 6
44	796.8	57.2	3754	6	A42478	A42478 Sequence 6
45	796.8	57.2	3754	6	A42478	A42478 Sequence 6

ALIGNMENTS

```

RESULT 1
LOCUS CLOTEFTOX
DEFINITION C.tetani tetanus toxin gene.
ACCESSION M12739.1 GI:144920
VERSION M12739.1 GI:144920
KEYWORDS tetanus toxin.
SOURCE C.tetani (strain CN3911, derivative of Harvard strain) DNA, clones pHE11.81.
ORGANISM Clostridium tetani
Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE 1 (bases 1 to 1791)
Fairweather,N.F., Lyness,V.A., Pickard,D.J., Allen,G. and
Thomson,R.O.

```

TITLE Cloning, nucleotide sequencing, and expression of tetanus toxin fragment C in *Escherichia coli*
 JOURNAL J. Bacteriol. 165 (1), 21-27 (1986)
 MEDLINE 86085672
 PUBMED 3510187

Draft entry and sequence in computer readable form for [1] kindly provided by N.F. Fairweather, 05-AUG-1986.
 The last 122 amino acids of fragment B are found at positions 1 to 366. Fragment C of the tetanus toxin is located at positions 367-1719.

FEATURES

source

CDS

Location/Qualifiers
 1..1791
 /organism="Clostridium tetani"
 /db_xref="taxon:1513"
 <1..1722
 /note="tetanus toxin"
 /codon_start=1
 /transl_table=1
 /protein_id="AA023282.1"
 /db_xref="GI:144921"
 /translation="RSLEYOVDAIKRLIDYEYKISGPRDEQIADETINLKNLEKKA
 NKAMININIMPSRSSRSLVNMOMEAKKOLLEDTOSKINLMQYIKANSKFLITEL
 KLESKINIKVPSPIPSYSKNIDCWDNEDIDILKSTILNDIINDIISDGF
 NSSVITYPDQOLVPIGINKAHLVNNSSSVYVHKAMDIEYNDMENNPTSEMLRVPK
 VSAHLEOYGTNEYSIISSMKHSLSTGSGSVSLKGNLITWTKDSAGEVROITPRD
 LPDFNAYLANRWEIFITNDRLSSANDLYINGVLMSEITGLGALREDNNTTKLDR
 CNNNOYVSLDKFIEFCALNPKEIKLYSYLSITELRDGWPMLRDEYLIPIVA
 SSSKVDOLKNTIDYVRLNAPSYNGRIKLYRNGPKIIRKYPNNIDSEFVS
 GDFIKLYVNNNHIVGPKDGNAPNNLDRIILVGYNAPPIPKKMEAVKLDLKT
 YSVOLKLYDXKNASLGLVGHINGQIGNDPNNDILIASWVYNHLKDKILGCDWYFVPT
 DEGWIND

BASE COUNT 727 a 197 c 276 g 591 t
 ORIGIN 1 bp upstream of BglII site.

Query Match 99.7% Score 1387.4; DB 1; Length 1791;
 Best Local Similarity 99.9%; Pred. No. 3.2e-174;
 Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTTTTTCACACCAATTCATTTCTTATCTAAATCTGATGTTGGGTGATAT 63
 DB 334 GTTTTTCACACCAATTCATTTCTTATCTAAATCTGATGTTGGGTGATAT 393
 QY 64 GAAGAAGATATGATGTTATTTAAAGAGTCAATTTTAAATTTAGATTTAATAT 123
 DB 394 GAAGAAGATATGATGTTATTTAAAGAGTCAATTTTAAATTTAGATTTAATAT 453
 QY 124 GATATTATATCAGATATATCTGGGTTAATTCATCTGTATATCAATCCAGTGTCAA 183
 DB 454 GATATTATATCAGATATATCTGGGTTAATTCATCTGTATATCAATCCAGTGTCAA 513
 QY 184 TTGGTCCCGGATTAATGGCAAGCAATATCATTTAGTAACAATGATCTTGAAGT 243
 DB 514 TTGGTCCCGGATTAATGGCAAGCAATATCATTTAGTAACAATGATCTTGAAGT 573
 QY 244 ATAGTCATTAAGCTATGATATTTGAATATTAAGATATGTTTAAATTTTACCGTTAGC 303
 DB 574 ATAGTCATTAAGCTATGATATTTGAATATTAAGATATGTTTAAATTTTACCGTTAGC 633
 QY 304 TTTTGGTTGAGGGTCCCTAAAGTATCTGCTAGTCATTTAACAATATGGCACAAATGAG 363
 DB 634 TTTTGGTTGAGGGTCCCTAAAGTATCTGCTAGTCATTTAACAATATGGCACAAATGAG 693
 QY 364 TATTCATATTTAGCTCTATGAAAAAACAATGATATCATATAGATGATCGTTGGAGTGA 423
 DB 694 TATTCATATTTAGCTCTATGAAAAAACAATGATATCATATAGATGATCGTTGGAGTGA 753
 QY 424 TCACCTTAAGGTAATACTTAATATGAGCTTTAAAGATTCGCGGAGAGAGTTAGCAA 483
 DB 754 TCACCTTAAGGTAATACTTAATATGAGCTTTAAAGATTCGCGGAGAGAGTTAGCAA 813
 QY 484 ATAACTTTAGGATTTACCGATTAATTTAATGCTTATTTAGCAAAATTAATGGGTTTTT 543
 DB 814 ATAACTTTAGGATTTACCGATTAATTTAATGCTTATTTAGCAAAATTAATGGGTTTTT 873

QY 544 ATAACTTTAGGATTTACCGATTAATTTAATGCTTATTTAGCAAAATTAATGGGTTTTT 603
 DB 874 ATAACTTTAGGATTTACCGATTAATTTAATGCTTATTTAGCAAAATTAATGGGTTTTT 933
 QY 604 GGAAGTCAGAAATTTAGCTGTTTGGAGCTATTTAGAGAGCTAATTAATTAATTAATAA 663
 DB 934 GGAAGTCAGAAATTTAGCTGTTTGGAGCTATTTAGAGAGCTAATTAATTAATTAATAA 993
 QY 664 CTAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
 DB 994 CTAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1053
 QY 724 AAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
 DB 1054 AAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1113
 QY 784 TTAAGAGACTTCTGGGGAACCCCTTACGATATGATACAGATATTAATTAATTAATTAAT 843
 DB 1114 TTAAGAGACTTCTGGGGAACCCCTTACGATATGATACAGATATTAATTAATTAATTAAT 1173
 QY 844 GCTTCTAGTTCTAAAGATGTTCAATTTGAATAATTAACAGATTAATTAATTAATTAAT 903
 DB 1174 GCTTCTAGTTCTAAAGATGTTCAATTTGAATAATTAACAGATTAATTAATTAATTAAT 1233
 QY 904 GCGGCATGCTATACAAAGAAATTTGAATATTAATTAATTAATTAATTAATTAATTAAT 963
 DB 1234 GCGGCATGCTATACAAAGAAATTTGAATATTAATTAATTAATTAATTAATTAATTAAT 1293
 QY 964 AAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
 DB 1294 AAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1353
 QY 1024 GATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
 DB 1354 GATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1413
 QY 1084 GATGGAATGCTTAAATATCTTGTAGTAATTTCTAAGATAGGTTATTAATGGCCCAAGGT 1143
 DB 1414 GATGGAATGCTTAAATATCTTGTAGTAATTTCTAAGATAGGTTATTAATGGCCCAAGGT 1473
 QY 1144 ATCCCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1203
 DB 1474 ATCCCTCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1533
 QY 1204 CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
 DB 1534 CAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1593
 QY 1264 CAATATGCGACAGATCCAAATAGGATATTTAATTCGAAGCAACTGTACTTTAATCAT 1323
 DB 1594 CAATATGCGACAGATCCAAATAGGATATTTAATTCGAAGCAACTGTACTTTAATCAT 1653
 QY 1324 TTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1383
 DB 1654 TTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1713
 QY 1384 AATGATTAA 1392
 DB 1714 AATGATTAA 1722

RESULT 2
 E01135 1851 bp DNA linear PAT 29-SEP-1997
 LOCUS E01135
 DEFINITION cDNA encoding tetanus toxin C fragment.
 ACCESSION E01135
 VERSION E01135.1 GI:2169394
 KEYWORDS JP 1987051994-A/1
 SOURCE Clostridium tetani.
 ORGANISM Clostridium tetani.
 Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

[illegible]

REFERENCE 1 (bases 1 to 4338)
 Fairweather, N.F. and Lyness, V.A.
 The complete nucleotide sequence of tetanus toxin
 JOURNAL Nucleic Acids Res. 14 (19), 7809-7812 (1986)
 MEDLINE 87040747
 PUBMED 3774547
 COMMENT Data kindly reviewed (25-Nov-1988) by Lyness V.A.
 FEATURES
 Location/Qualifiers
 1. 4338
 /organism="Clostridium tetani"
 /strain="CN3911"
 /db_xref="taxon:1513"
 /clone="PAT153"
 322..4269
 /note="tetanus toxin (AA 1-1315)"
 /codon_start=1
 /protein_table=11
 /protein_id="CAA29564.1"
 /db_xref="GI:40774"
 /db_xref="SWISS-PROT:P04958"
 /translation="MPITINFRYSDPVNNDTIIMPEPYCKGLDIYKAFKLTIDRIM
 IYPEREFGKPEDENPSLIIGASEYDNPVLPDSDKDFLOTMVKLFNKKNV
 ACGLLDKTIINAIPYGSYSLIDKEDTMSYSEFLIDODSGATTKSAMLTILF
 GPEVLNKEVKGIVLRDNKNTFPCRDGSGSIMQAFCEPYVPEPDNTENTSLTI
 GSKRYQDPALMLHELIVLHGLYGMVSSHETPSKDEITVQHTPIISAEELPFG
 GDANLISIDIKNDLYEKTLDYKALANKLSQVTSNDPDIIDISTKQYIOQVYFDK
 DSNQGYIVNEDKFOILYNSIMGFTEIELGKFNITRLSYFNMNDPVKIPMLDPT
 IYDTEGPNIESKDLSEYKQGMRYNAFRVDSGIVSKLIGLCKKIIPNTIRE
 NIYNRTASLIDIGGELCIKIKNEDLFTIEAKNSEEPDEDEIVSNTRKPLNFVS
 LDKITDYNLQSKITLPNDPTPTPKGIGPAPRYKNAASTIEIHNDIOWROITDPT
 OKSPFTLQRTITNSVDALINSTKIYSTPYSISKVSAQGITLLELOWROITDPT
 NNESSOKTYYDGIISDVIVYIGPALNIVQSGEYNGALLETGVVLEIETPEPIL
 PVTAALSIESTQKEKIIKTIDNPEKREKWEYKIVKAKWLGKVNQFQRSTQ
 MRSLEOVDAIKKIIDYEKKIYSGPDEQIAEINLKNKLEKANKAMININIMR
 ESSRSFLYNOMINEAKKOLEPDTOSKILMOYIKANSKFIGITELKLESKINKVS
 TPPIPEYKSKLDQVNDNEIDIVLIKSTLINDINDIISDISGESSVITYPDOL
 VPEINGKAIHLVNNESSEYVYHKAAMDIEYDMNPNFVSFLIEPVYSAHLIOYGN
 EYSIISSMKHSLISGSMSVLKGNLITLTDOSAGEVQITPRDIPKFNALANK
 WVEITITNDRLSASNLINGVLMSAETIGLAIREDNITLTLDRNNNOYVLDK
 FRFEKALNKEIEKELNLYSYSLTFLEPDEGNLRDTEYLLIPVASSSDVOLKNT
 DYVLTNAPSYTNGKLNLYYRRLNGKLEIKRYTPNNEIDSEVKSDFIKLVSYNN
 NEHIVGPKDGNAPNNLDRIIRYGVNAPGILVKKMEVAILRDLTKTSYOLKLYDDKN
 ASLIGVTHNGOIGDNPNDILLIASNMYFHLMDKILGCMWYFVPTDEGMTND"

BASE COUNT 1783 a 522 c 640 g 1393 t
 ORIGIN
 Query Match 99.7%; Score 1387.4; DB 1; Length 4338;
 Best Local Similarity 99.9%; Pred. No. 2.4e-174;
 Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTTTTTCACACCAATTCATTTCTTATTCCTAAATATCGGATTTGGTGGTATAT 63
 |||||||
 Db 2881 GTTTTTCACACCAATTCATTTCTTATTCCTAAATATCGGATTTGGTGGTATAT 2940
 |||||||
 QY 64 GAAGAGATATAGATATATTAATAAAGAGTACATTTTAAATTTGATATTAAT 123
 |||||||
 Db 2941 GAAGAGATATAGATATATTAATAAAGAGTACATTTTAAATTTGATATTAAT 3000
 |||||||
 QY 124 GATATTATATCAGATATATCTGGGTTTAATTCATCTGTAATAACATATCCAGATCTCAA 183
 |||||||
 Db 3001 GATATTATATCAGATATATCTGGGTTTAATTCATCTGTAATAACATATCCAGATCTCAA 3060
 |||||||
 QY 184 TTGGTGCCCGGATAAAGCGCAACATACATTAGTAACAATAAGATCTTCCGAAGTT 243
 |||||||
 Db 3061 TTGGTGCCCGGATAAAGCGCAACATACATTAGTAACAATAAGATCTTCCGAAGTT 3120
 |||||||
 QY 244 ATAGTGCATTAAGTATGATATTAATTAATGATATGATATTAATTTACCGTTAGC 303
 |||||||
 Db 3121 ATAGTGCATTAAGTATGATATTAATTAATGATATGATATTAATTTACCGTTAGC 3180
 |||||||
 QY 304 TTTTGTTGAGGGTTCCTAAGTATCTGCTAGTCAATTAAGAACATATGGCAATGAG 363
 |||||||

Db 3181 TTTTGTGAGGTTCTCCTAAGATATCTGCTAGTCAATTAAGAACATAATGCGACAAATGAG 3240
 |||||||
 QY 364 TATTCATTAATTAAGTCTCTATGAAGAAAAACATAGCTCTATCATATAGGATCTGGTGAGGTGA 423
 |||||||
 Db 3241 TATTCATTAATTAAGTCTCTATGAAGAAAAACATAGCTCTATCATATAGGATCTGGTGAGGTGA 3300
 |||||||
 QY 424 TCACCTTAAGGATATACCTTAATATGAGCTTTAAAGATCCCGGAGAGTATAGCAA 483
 |||||||
 Db 3301 TCACCTTAAGGATATACCTTAATATGAGCTTTAAAGATCCCGGAGAGTATAGCAA 3360
 |||||||
 QY 484 ATACCTTTTATAGGATTTACCTGATAAATTTAATCTTATTTAGCAATAAATGGGTTTTT 543
 |||||||
 Db 3361 ATACCTTTTATAGGATTTACCTGATAAATTTAATCTTATTTAGCAATAAATGGGTTTTT 3420
 |||||||
 QY 544 ATACCTTTTATAGGATTTACCTGATAAATTTAATCTTATTTAGCAATAAATGGGTTTTT 603
 |||||||
 Db 3421 ATACCTTTTATAGGATTTACCTGATAAATTTAATCTTATTTAGCAATAAATGGGTTTTT 3480
 |||||||
 QY 604 GGAAGTGCAGAAATTTACTGGTTTATAGAGCTATTATAGAGGATTAATAATATATATATAA 663
 |||||||
 Db 3481 GGAAGTGCAGAAATTTACTGGTTTATAGAGCTATTATAGAGGATTAATAATATATATAA 3540
 |||||||
 QY 664 CTAGATAGATGTAATAATTAATTAATCAATACGTTCTATTTGATAATTTAGATATTTTGC 723
 |||||||
 Db 3541 CTAGATAGATGTAATAATTAATTAATCAATACGTTCTATTTGATAATTTAGATATTTTGC 3600
 |||||||
 QY 724 AAGCATTAATCCCAAGAGATTTGAAGAAATTTATACCAAGTTATTAATCTATACCTTT 783
 |||||||
 Db 3601 AAGCATTAATCCCAAGAGATTTGAAGAAATTTATACCAAGTTATTAATCTATACCTTT 3660
 |||||||
 QY 784 TTAAGAGACTTCTGGGGAACCCCTTTACGATATGATACAGAAATTTATTAATACCGATA 843
 |||||||
 Db 3661 TTAAGAGACTTCTGGGGAACCCCTTTACGATATGATACAGAAATTTATTAATACCGATA 3720
 |||||||
 QY 844 GCTCTAGTCTTAAAGATGTTCAATTTGAAAATATTAACAGATTTATAGTATTTTGCAAT 903
 |||||||
 Db 3721 GCTCTAGTCTTAAAGATGTTCAATTTGAAAATATTAACAGATTTATAGTATTTTGCAAT 3780
 |||||||
 QY 904 GCGGCATGCTATCTAACGGAATTTGAATATTAATTTAAGAGTTATTAATAGGACTA 963
 |||||||
 Db 3781 GCGGCATGCTATCTAACGGAATTTGAATATTAATTTAAGAGTTATTAATAGGACTA 3840
 |||||||
 QY 964 AAATTTATTAATAAAGATATACCTTAATTAATGAATATAGATTTCTTTGTTAATACAGT 1023
 |||||||
 Db 3841 AAATTTATTAATAAAGATATACCTTAATTAATGAATATAGATTTCTTTGTTAATACAGT 3900
 |||||||
 QY 1024 GATTTTATTAATATATATGATATATATATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
 |||||||
 Db 3901 GATTTTATTAATATATATGATATATATATTAATTAATTAATTAATTAATTAATTAATTAAT 3960
 |||||||
 QY 1084 GATGGAATATGCTTAAATATATCTTGAATAGAAATTTAAGAGTAGGTTATTAATGCCAGGT 1143
 |||||||
 Db 3961 GATGGAATATGCTTAAATATATCTTGAATAGAAATTTAAGAGTAGGTTATTAATGCCAGGT 4020
 |||||||
 QY 1144 ATCCCTCTTTATTAATAAATATGACAGTAAATTTGCGTGAATTTAAACCAATCTGTGA 1203
 |||||||
 Db 4021 ATCCCTCTTTATTAATAAATATGACAGTAAATTTGCGTGAATTTAAACCAATCTGTGA 4080
 |||||||
 QY 1204 CAACCTTAATTAATATGATATGATTAATAATGATCTTTAGGACTAGTAGTACCCATATGCT 1263
 |||||||
 Db 4081 CAACCTTAATTAATATGATATGATTAATAATGATCTTTAGGACTAGTAGTACCCATATGCT 4140
 |||||||
 QY 1264 CAATATAGCAACGATCCAAATAGGATATATTAATTAATGACAGCACTGTTATTAATCAT 1323
 |||||||
 Db 4141 CAATATAGCAACGATCCAAATAGGATATATTAATTAATGACAGCACTGTTATTAATCAT 4200
 |||||||
 QY 1324 TTTAAAGATTAATTTTATAGATATGATATGATATGATATGATATGATATGATATGATATGAT 1383
 |||||||
 Db 4201 TTTAAAGATTAATTTTATAGATATGATATGATATGATATGATATGATATGATATGATATGAT 4260
 |||||||
 QY 1384 AATGATTAAT 1392
 |||||||
 Db 4261 AATGATTAAT 4269
 |||||||

RESULT 4
LOCUS CTTETANG 4400 bp DNA linear BCT 12-JUL-1995
DEFINITION Clostridium tetani gene for tetanus toxin.
ACCESSION X04436
VERSION X04436.1 GI:40769
KEYWORDS neurotoxin; plasmid; tetanus toxin; toxin.
SOURCE Clostridium tetani.
ORGANISM Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.
REFERENCE 1 (bases 1 to 4400)
Eisel, U., Jarausch, W., Goretzki, K., Henschen, A., Engels, J., Weller, U., Hidel, M., Habermann, E. and Niemann, H.
Tetanus toxin: primary structure, expression in E. coli, and homology with botulinum toxins
EMBO J. 5 (10), 2495-2502 (1986)
JOURNAL MEDLINE 87053814
COMMENT PUBMED 3536478
See also Fairweather et al. (1986) J. Bacteriol. 165, 21-27. Data kindly reviewed (12-MAY-1987) by U. Eisel.
FEATURES
source
1. 4400
/organism="Clostridium tetani"
/db_xref="taxon:1513"
<1..50
/note="Open reading frame (16AA) (1 is 2nd base in codon)"
/codon_start=1
/transl_table=1
/protein_id="CAA28032.1"
/db_xref="GI:809667"
/db_xref="SPTREMBL:Q46456"
/translation="ALKLEPIVKNLINI"
37..250
/note="region of inverted repeats"
267..272
/note="rRNA binding site"
281..4228
/codon_start=1
/transl_table=1
/product="tetanus toxin precursor (AA 1-1315)"
/protein_id="CAA28033.1"
/db_xref="GI:40770"
/db_xref="SWISS-PROT:P04958"
/translation="MPTINNFRYSDEPVNNDIIMMEPYCKGLDIYYKAEKITDRIM IVERLYEFGTKPEDFNPSSLIEGASEYDNYLRTDSDKRFLOTWKLFEKRIKNV AGEALDKIINALPYLGNYSLLDKPDPNSVSFNLEOPSGATTSAMITNLITF GPGVILKNENRGIVLRVKNKYPFCRDRGCSIMOMAFCPREVPEDVIERITSLTI GKSKYFDDPALLMHELIVHLHGIXGMQVSHETIPSKOETIMORTYISAELETFG GQANLISIDIKNDLYEKTLLNDYKALANKLSQVTSNDPNIDISYKQIYQKQYFDK DSKQYIVNEDKFDILNYSIMYGFTELEGKFNKIRLSYFSMHDVVKIIPNLDDT IYNDTEGFNESDKSEYKGNMKNVNTNFAENVDSGLVGLGCKIIPNLIRE NLNRTASLIDLGELCIKIKNEDLTFIAEKNSFESEFOEIVSNKKNIPNINYS LDKIIVYNGOSKLTLPNDRTTPYTKGIPVAPAEKSNASATLIEIHNIDNNIYQVLYA OKSPTORTITMNSVDALINSTKIYSPSVISKVNOGAGLIFLWDNDITDFT NESOKRTDKISVSITVYIPGALNTVKGYSNGEFGALETGVLVLEWDITDFT PVTAAIATASSTOKEKIKITIDNLEKREKLEVKLVAKMGLVNTQFORSKO MYRSLEYOVAIKRIIDYEKIKYSGPKDEQLADEINNKLEKANKAMININEMR ESSRSFLVNMINEAKROLEPDTOSKNILMOYIKANSKFTIGTELKLESKINVS TPIPFYSKNDICVDNEEDIDELTKSTLILNDINNDISIDGPNSSVITYPACOL VPGNGKAILHVNNESEVIVKAMDIENYDMFNPNFSPVPLRVPKVSASHEBOGTN EYSILSMKHSLSIGSGMSVSLKGNLILITGAIREDNNITLKLIDRCNNNNQVSIADK KWTITINDLSANLITNGVLMGSALITGAIREDNNITLKLIDRCNNNNQVSIADK FRIEFCALNKEIEKLYTSYSTLTFLDFMGNPLRYDEYVLLIVASSKDVQAKNIT DYMVLINAPSYTNGKLNLYRRLNGLKFTIKRYTPNNEIDSEVSKSGDFIKLYSYNN NEHIVGPKDGNAPFNIDRLIRVGNAPGLPIYKMEAVKLRDCKTYSVOLKLVDDKN ASLIGVTHNGQICNDPNRIDILASNMVFHMKKIKIIGCDWYFVPTDEGTND"
281..1651
/product="tetanus toxin L-chain (AA 1-457)"
1652..4225
/product="tetanus toxin H-chain (AA 458-1315)"
mat_peptide
mat_peptide

misc_feature 4268..4394
/note="region of inverted repeats"
BASE COUNT 1808 a 532 c 644 g 1416 t
ORIGIN
Query Match 99.7%; Score 1387.4; DB 1; Length 4400;
Best Local Similarity 99.9%; Pred. No. 2.4e-174;
Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 GTTTTTCACACCAATTCATTTCTTATTTCTTAAAAATCTGGATTTGGGTGATTAAT 63
|||||
Db 2840 GTTTTTCACACCAATTCATTTCTTATTTCTTAAAAATCTGGATTTGGGTGATTAAT 2899
QY 64 GAAGAGATATAGAT 123
|||||
Db 2900 GAAGAGATATAGAT 2959
QY 124 GAT 183
|||||
Db 2960 GAT 3019
QY 184 TTGGTGGCCGGAATTAAGGCAAGACATATATATATATATATATATATATATATATAT 243
|||||
Db 3020 TTGGTGGCCGGAATTAAGGCAAGACATATATATATATATATATATATATATATATAT 3079
QY 244 ATAGTGAATTAAGTAT 303
|||||
Db 3080 ATAGTGAATTAAGTAT 3139
QY 304 TTTTGTGTTGAGGTTTCTTAAGATATCTGCTAGTCAATTAAGCAATATGCGCAATATGAG 363
|||||
Db 3140 TTTTGTGTTGAGGTTTCTTAAGATATCTGCTAGTCAATTAAGCAATATGCGCAATATGAG 3199
QY 364 TATTCATTAATTTACCTCTATGAAAAAACAATAGTATATATATATATATATATATATAT 423
|||||
Db 3200 TATTCATTAATTTACCTCTATGAAAAAACAATAGTATATATATATATATATATATATAT 3259
QY 424 TCACCTTAAGGTAATTAAGTAT 483
|||||
Db 3260 TCACCTTAAGGTAATTAAGTAT 3319
QY 484 ATACCTTTTGGGATTTACCTGATTAATTTATATATATATATATATATATATATATATAT 543
|||||
Db 3320 ATACCTTTTGGGATTTACCTGATTAATTTATATATATATATATATATATATATATATAT 3379
QY 544 ATACCTTTTACTAT 603
|||||
Db 3380 ATACCTTTTACTAT 3439
QY 604 GGAAGTGCAGAAATTTACTGTTTGAAGCTATTAAGAGATTAATTAATTAACATTAATAA 663
|||||
Db 3440 GGAAGTGCAGAAATTTACTGTTTGAAGCTATTAAGAGATTAATTAATTAACATTAATAA 3499
QY 664 CTGAT 723
|||||
Db 3500 CTGAT 3559
QY 724 AAAGCATTAATCAAAAGAGATTAAGAAATTTATACCAATTTATTTCTATATACCTTT 783
|||||
Db 3560 AAAGCATTAATCAAAAGAGATTTGAAAATTTATACCAATTTATTTATCTATATACCTTT 3619
QY 784 TTAAGACATTTGGGGAAACCTTTACGATATGATACAGAAATTTATTTATTAATACAGTA 843
|||||
Db 3620 TTAAGACATTTGGGGAAACCTTTACGATATGATACAGAAATTTATTTATTAATACAGTA 3679
QY 844 GCTTTAGTCTTAAAGTGTTCAAATGAATAATTTATACAGATTTATATGTTATTGACAAAT 903
|||||
Db 3680 GCTTTAGTCTTAAAGTGTTCAAATGAATAATTTATACAGATTTATATGTTATTGACAAAT 3739
QY 904 GCGGCATGCTATATTAAGGAAATTAATATATATATATATATATATATATATATATATAT 963
|||||
Db 3740 GCGGCATGCTATATTAAGGAAATTAATATATATATATATATATATATATATATATATAT 3799

OY	964	AAATTTTPTTTTTAAAAACATATATACCCCAATAAAGCAATPAGCTTCCTTTGGTTAAATCAGGT	1023
Db	3800	AAAATTATTATTAATAAAGAATATACACCTPAATPATGAAATAGACTTCCTTTTGTTAAATCAGGT	3859
OY	1024	GATTTTATTTAAATATATATGTATCATATATAACAATTAATAGAGCACATTTGATAGCTTATCCGAAA	1083
Db	3860	GATTTTATTTAAATATATATNGTATCATATATAACAATTAATAGAGCACATTTGATAGCTTATCCGAAA	3919
OY	1084	GATGGAATATGCCCTTTAAATATCTTGATAGAAATCTTAAGACTAGGTATATGCCAGGT	1143
Db	3920	GATGGAATATGCCCTTTAAATATCTTGATAGAAATCTTAAGAGAGGTATTAATGCCCCAGGT	3979
OY	1144	ATCCCTCTTTTAAAAAAGTAGAGACAGTAAATTCGGTGTATTTAAAAACCCTATCTGTGA	1203
Db	3980	ATCCCTCTTTTAAAAAAGTAGAGACAGTAAATTCGGTGTATTTAAAAACCCTATCTGTGA	4039
OY	1204	CACCTTAATTTATATATGATGATGATAAAAATGCATCTTTAGAGACTAGTAGCTAACCATPATGCT	1263
Db	4040	CACCTTAATTTATATATGATGATGATAAAAATGCATCTTTAGAGACTAGTAGCTAACCATPATGCT	4099
OY	1264	CAATATAGCAACCGATCCCAATATAGGATATATTAATTTGGAACCACTGGTACTTTAAATCAT	1323
Db	4100	CAATATAGCAACCGATCCCAATATAGGATATATTAATTTGGAACCACTGGTACTTTAAATCAT	4159
OY	1324	TTAAAAAGATAAAAATTTTGGATGTGATGTGACTCTTTGACCTACAGATGAGAGATGAGACA	1383
Db	4160	TTAAAAAGATAAAAATTTTGGATGTGATGTGACTCTTTGACCTACAGATGAGAGATGAGACA	4219
OY	1384	AATGATTAA 1392 	
Db	4220	AATGATTAA 4228	
RESULT 5			
LOCUS	AR017561	1858 bp	DNA linear PAT 05-DEC-1998
DEFINITION	Sequence 6 from patent US 5780024.		
ACCESSION	AR017561		
VERSION	AR017561.1	GI:3973164	
KEYWORDS			
SOURCE	Unknown. Unclassified.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1858)		
AUTHORS	Brown,R.H., Fishman,P.S., Francis,J.W. and Hosler,B.A.		
TITLE	Superoxide dismutase/tetanus toxin fragment C hybrid protein		
JOURNAL	Patent: US 5780024-A 6 14-JUL-1998;		
FEATURES	Location/Qualifiers source 1..1858 /oranism="unknown"		
BASE COUNT	665 a 248 c 373 g 572 t		
ORIGIN			
Query Match	99.1%;	Score 1379.8;	DB 6;
Best Local Similarity	99.9%;	Pred. No. 3.2e-173;	Length 1858;
Matches 1381; Conservative	0;	Mismatches 2;	Indels 0;
			Gaps 0;
OY	10	TCACACCACAATTCACATTTCTTATCTTAAATAATCTGATTTGGGTGATATGAGAA	69
Db	476	TCACACCACTTCCATTTCTTATCTTAAATAATCTGATTTGGGTGATATGAGAA	535
OY	70	GATATAGATGTTATATTAATAAAGAATACAAATTTAAATTTAGATATTAATATGATATT	129
Db	536	GATATAGATGTTATATTAATAAAGAATACAAATTTAAATTTAGATATTAATATGATATT	595
OY	130	ATATCAGATATATCTGGGTTTAATTCATCTGTAATATACATATCCAGATGCTCAATTGGTG	189
Db	596	ATATCAGATATATCTGGGTTTAATTCATCTGTAATATACATATCCAGATGCTCAATTGGTG	655
OY	190	CCCAGATTAATAGCAACAACATATATAGTAATACAAATGATCTTCTCAAGTTATAGTG	249
Db	656	CCCAGATTAATAGCAACAACATATATAGTAATACAAATGATCTTCTCAAGTTATAGTG	715

QY	250	CATAAGCTATGAGTAATTGAAATATATACATATAGTTAAATTAATTTACCGTTGACCTTTTGG	309
Db	716	CATAAAGCTATGCAATTTGGAATATATAGTATATGTTTAATTAATTTTACCGTTTACGCTTTTGG	775
QY	310	TTGAGGGTTCCAAAGATATCTGCTAGTACTATTTAGAACATATATGGACAACAATGAGTATTTCA	369
Db	776	TTGAGGGTTCCAAAGATATCTGCTAGTACTTTAGAACATATATGGACAACAATGAGTATTTCA	835
QY	370	ATAATTTAGCTATGAAAAAACATATGCTATCATATATAGATCTTGTTGGAGGTATACACT	429
Db	836	ATAATTTAGCTATGAAAAAACATATGCTATCATATAGATCTTGTTGGAGGTATATACACT	895
QY	430	AAAGCAATTAACCTTATATATGACCTTTAAAGAATTCGCGGAGAGATTTAGACAAATTAAC	489
Db	896	AAAGCAATTAACCTTATATATGACCTTTAAAGAATTCGCGGAGAGATTTAGACAAATTAAC	955
QY	490	TTTTAGGGATTTTACCTGATTAATTTAAATGCTTATTTAGCAATTAATGGGTTTTTATTAAC	549
Db	956	TTTTAGGGATTTTACCTGATTAATTTAAATGCTTATTTAGCAATTAATGGGTTTTTATTAAC	1015
QY	550	ATTACTAATGATAGATTAATTAATCTGCTGATTTGTTATTAATGAGACTACTTATGGGAAC	609
Db	1016	ATTACTAATGATAGATTAATTAATCTGCTGATTTGTTATTAATGAGACTACTTATGGGAAC	1075
QY	610	GCAGAAATTTACTGTTTAAAGACCTATTAGAGAGGATATATTAACATTTAAACTAGAT	669
Db	1076	GCAGAAATTTACTGTTTAAAGACCTATTAGAGAGGATATATTAACATTTAAACTAGAT	1135
QY	670	ACATGTAATTAATTAATTAATCAATACGTTCTTATTTGATTAATTTAGATTTATTTGGCAANGA	729
Db	1136	ACATGTAATTAATTAATTAATCAATACGTTCTTATTTGATTAATTTAGATTTATTTGGCAANGA	1195
QY	730	TTAAATCCCAAAAGAGATTGAAAAATATACACAAGTTATTTATCTATTAACCTTTTAA	789
Db	1196	TTAAATCCCAAAAGAGATTGAAAAATATACACAAGTTATTTATCTATTAACCTTTTAA	1255
QY	790	GACTTCTGGGAAACCCCTTTACGATATGATACAGATATTTATTTAATACCAGTAGCTTCT	849
Db	1256	GACTTCTGGGAAACCCCTTTACGATATGATACAGATATTTATTTAATACCAGTAGCTTCT	1315
QY	850	AGCTTTAAAGATTTTCATTTGAAAAATTAACAGATTAATATGATTTTGAACAAATGGCCCA	909
Db	1316	AGCTTTAAAGATTTTCATTTGAAAAATTAACAGATTAATATGATTTTGAACAAATGGCCCA	1375
QY	910	TGCTATACTAACGGAAATTTGAATATATATTTATAGAAAGTTTATATTAATGGACATAAATTT	969
Db	1376	TGCTATACTAACGGAAATTTGAATATATATTTATAGAAAGTTTATATTAATGGACATAAATTT	1435
QY	970	ATTATTAATAAAGATATACACCTTAATTAATGAAATGATTTCTTTGTAAATAGCTGATTTT	1029
Db	1436	ATTATTAATAAAGATATACACCTTAATTAATGAAATGATTTCTTTGTAAATAGCTGATTTT	1495
QY	1030	ATTTAAATTTATATGATATCATATTAACAATAATGAGCACATTTCTAGCTTATCCGAAAGATGA	1089
Db	1496	ATTTAAATTTATATGATATCATATTAACAATAATGAGCACATTTCTAGCTTATCCGAAAGATGA	1555
QY	1090	AATGCCCTTAATTAATCTGTTGATTAATTTCTAAGAGTATAGGTATTAATGGCCAGGATATCC	1149
Db	1556	AATGCCCTTAATTAATCTGTTGATTAATTTCTAAGAGTATAGGTATTAATGGCCAGGATATCC	1615
QY	1150	CTTTATTAATAAATAATGAGACGATTAATAATTCGATTTTAAAAACCTATTTCTGTACAACTT	1209
Db	1616	CTTTATTAATAAATAATGAGACGATTAATAATTCGATTTTAAAAACCTATTTCTGTACAACTT	1675
QY	1210	AAATTTATATGATATATAAATATGATCTTTTGGAGCTATAGTACCAATATATGGCAATTA	1269
Db	1676	AAATTTATATGATATATAAATATGATCTTTTGGAGCTATAGTACCAATATATGGCAATTA	1735
QY	1270	GGACACGATCCAATATAGGAGATATATTAATTTATGCAAGCAACAATCGTACTTTTAATCAATTAA	1329
Db	1736	GGACACGATCCAATATAGGAGATATATTAATTTATGCAAGCAACAATCGTACTTTTAATCAATTAA	1795
QY	1330	GATTAATAATTTTACGATGTGATTGGTACTTTGTACTCTACAGATGAGGATGAGCAAAATGAT	1389

|||||
Db 1796 GATTAATTTAGATGTGTGTTACTTTGTACTACCTACAGATGAAGATGACCAATGAT 1855
OY 1390 TAA 1392
|||
Db 1856 TAA 1858
RESULT 6
AF389424 3931 bp DNA linear BCT 26-JUL-2001
LOCUS Clostridium tetani tetanus toxin gene, partial cds.
DEFINITION AF389424
ACCESSION AF389424
VERSION AF389424.2 GI:15022164
KEYWORDS
SOURCE Clostridium tetani.
Clostridium tetani.
ORGANISM Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
Clostridium.
REFERENCE 1 (bases 1 to 3931)
AUTHORS Shumin,Z. and Dianliang,L.
TITLE Cloning and sequence analysis of tetanus toxin gene
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3931)
AUTHORS Shumin,Z. and Dianliang,L.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2001) Department of Serum, National Institute for
the Control of Pharmaceutical and Biological Products, No. 2
Tiantanxili, Temple of Heaven, Beijing 100050, P.R. China
COMMENT On Jul 26, 2001 this sequence version replaced by:14718605.
FEATURES
source location/Qualifiers
1..3931
/organism="Clostridium tetani"
/db_xref="taxon:1513"
<1..>3931
/codon_start=2
/transl_table=11
/product="tetanus toxin"
/protein_id="AAK72964.2"
/db_xref="GI:15022165"
/translation="RIPPIINNRYSVPVNNPTIIMPPRYCKGLDYYKAKFTTORT
WVPERYEGTKPEDPNPSSSIEGASEYIDENYLTDSKDFLTQWKLFRIRKN
VAGEALDKITNAIPPLGNSYSLDEFDNNSVSNLEODSGATTSKAMLSLI
FGPGVLKNEVRGIVLRVDNKNRPPCRGFGSINMTCPEPFDNYENITSLT
IGSKRYFDPDALLMHEILHVLHGLYGMVSSHEIIPSKOEIYMOHYPIASAELEF
GGODANLISIDIKNDLYEKTLDYKAIKAKISOVTSCNPNDIDSKOJYOXYOD
KDSNGOXYINEDKPOLIVNSIMYGETELGKKFKPTLSYSNMHPYKKTINLDD
TIYNDEGENIESKDLSEYKQONNRVNTNARNVDSGLVSKMLKIPPTNR
ENLIRKIALDLGELCLIKINEDLIFAEKNSESEEPQDEIVSYNKKPLNFNY
SLDKIILLYNLOSKITLPNDRTTPYKGIPIPAEYKSNASSTIEIHINDNTIYOYLY
AOKSPTLORTMTNSVDALINSTIYSPSVISKVNOGAGILFLOVRDIIIDE
TNESOKTITDKISDVSTIPIYIPALNTIKVGGEGNFIAGLETTGVILLIETP
LPVIALSTAESSTQKEKIKITIDNLEKRYEKMEIVYLKLVAKWLGTVNTOQKSY
QMRBLEYOVDAIKITIDYKITYSGPDEQLADEINLNKLNKEENANKAMININIFM
RESSSPFLVNMINETKQLEFDPQSKNIMQYIKANSFISITELKLESKINVE
STPIEFSYKNDICWDMEDIDVLKSTIINLDNIIISDISFNSVITYPAOL
LVPGINGKAIRLVNNSESVIYKAMDIIYNDMPNEFVSFMRVKSVAEYPOD
NEYSTISMKKYSLSIGSGVSIKGNLIMLKLDSAGVROTTFPDLSPKRNAYLAN
KWFETITNDRSSANLYINGVLMGSAETTGALIEDNNITLKLDRNNNOYVSD
KFRICKALNKEIEKLYTSYLTFLRFQWGPPLDYDEYLIIPAYSSSKDQLNKI
TDYMYLTNAPSTYNGKRLNLYRRLYSGLEFIIRKYPNNEIDISFVSGCFIKLYVSN
NNEHIVGYPKDNANFNNDRLIRVGNAGFIPLKMEAVKGLDLKTYVSQULKLYDOK
DASLGLVGHNGOIGNDPMRDILIASNMYFNHLKDKTLICDMVFVPTD"

64 GAAGAAGATATAGATGTATATTAATAAAGAGACATTTTAATTTAGATTAATAT 123
|||||
Db 2624 GAAGAAGATATAGATGTATATTAATAAAGAGACATTTTAATTTAGATTAATAT 2663
OY 124 GATATTATATCAGATATATCTGGGTTTAATTCATCTGTAAATTAACATATCCAGATGCTCAA 183
|||||
Db 2684 GATATTATATCAGATATATCTGGGTTTAATTCATCTGTAAATTAACATATCCAGATGCTCAA 2743
OY 184 TTGGTCCCGGAATTAATGCGCAAGCAATATACATTTGTAACATGAACTCTCGAAGTT 243
|||||
Db 2744 TTGGTCCCGGAATTAATGCGCAAGCAATATACATTTGTAACATGAACTCTCGAAGTT 2803
OY 244 ATAGTGCATTAAGCTATGATATATGATATTAATGATATGTTTATATTTTACCCTGAGC 303
|||||
Db 2804 ATAGTGCATTAAGCTATGATATATGATATTAATGATATGTTTATATTTTACCCTGAGC 2863
OY 304 TTTTGGTGGAGGGTTCCTTAAGTATCTGCTAGTCATTTAGAACATATATGCGACAAATGAG 363
|||||
Db 2864 TTTTGGTGGAGGGTTCCTTAAGTATCTGCTAGTCATTTAGAACATATATGACACAAATGAG 2923
OY 364 TATTCATATATATAGCTATGATTAAGAAAAACATAGCTATCATATAGATCTGGTGGAGTGA 423
|||||
Db 2924 TATTCATATATATAGCTATGATTAAGAAAAATATAGCTATCATATAGATCTGGTGGAGTGA 2983
OY 424 TCACCTTAAAGTAAATTAATTAATATATGATGACTTTAAAGATCCCGGAGAGATTAGACAA 483
|||||
Db 2984 TCACCTTAAAGTAAATTAATTAATATATGATGACTTTAAAGATCCCGGAGAGATTAGACAA 3043
OY 484 ATACTTTTGGAGATTTACCTGATATAATTTAATGCTTATTTAGCAATAAATGAGGTTTTT 543
|||||
Db 3044 ATACTTTTGGAGATTTATCTGATATAATTTAATGCTTATTTAGCAATAAATGAGGTTTTT 3103
OY 544 ATACTTATTAATTAATGATTAATTTCTGCTAATTTGATATTAATGAGTACTATG 603
|||||
Db 3104 ATACTTATTAATTAATGATTAATTTCTGCTAATTTGATATTAATGAGTACTATG 3163
OY 604 GGAATGCGAAGATTAATGATTAATTTCTGCTAATTTGATATTAATGAGTACTATG 663
|||||
Db 3164 GGAATGCGAAGATTAATGATTAATTTCTGCTAATTTGATATTAATGAGTACTATG 3223
OY 664 CTAGATAGATGATTAATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 723
|||||
Db 3224 CTAGATAGATGATTAATTAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 3283
OY 724 AAAGATTAATAACCAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 783
|||||
Db 3284 AAAGATTAATAACCAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 3343
OY 784 TTAAGAGACTTTGGGGAAACCTTTAGCATATGATAGCAATATTAATTAATACCGTA 843
|||||
Db 3344 TTAAGAGACTTTGGGGAAACCTTTAGCATATGATAGCAATATTAATTAATTAATACCGTA 3403
OY 844 GCTTCTAGTCTTAAGATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
|||||
Db 3404 GCTTCTAGTCTTAAGATGTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3463
OY 904 GCGCATGCTATACACGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
|||||
Db 3464 GCGCATGCTATACACGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3523
OY 964 AAATTTATTAATAAAGATATACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
|||||
Db 3524 AAATTTATTAATAAAGATATACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3583
OY 1024 GATTTTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
|||||
Db 3584 GATTTTATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3643
OY 1084 GATGGAATGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
|||||
Db 3644 GATGGAATGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3703

QY 1144 ATCCCTTTTATTAATAATGAGACAGTAAATTCGATGATTTAAATTAATTCCTGTA 1203
 |||||||
 Db 3704 ATCCCTCTTTTATTAATAATGAGACAGTAAATTCGATGATTTAAATTAATTCCTGTA 3763
 |||||||
 QY 1204 CAACCTTAATTAATGATGATTAATAATGATCTTTAGAGCTAGTAAAGTAAAGT 1263
 |||||||
 Db 3764 CAACCTTAATTAATGATGATTAATAATGATCTTTAGAGCTAGTAAAGTAAAGT 3823
 |||||||
 QY 1264 CAATAGGACAGTCCCAATAGGATATTAATGCAACCACTGCTACTTAATCAT 1323
 |||||||
 Db 3824 CAATAGGACAGTCCCAATAGGATATTAATGCAACCACTGCTACTTAATCAT 3883
 |||||||
 QY 1324 TTAAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
 |||||||
 Db 3884 TTAAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3931
 |||||||
 RESULT 7
 A17527 1359 bp DNA linear PAT 09-JUL-2002
 LOCUS human_firagment C tetanus gene seq ID No:1.
 DEFINITION A17527
 ACCESSION A17527
 VERSION A17527.1 GI:21727077
 KEYWORDS
 SOURCE human.
 ORGANISM human sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1359)
 AUTHORS McKoff,A.J., Romanos,M.A., Clare,J.J. and Fairweather,N.F.
 TITLE Vaccines
 JOURNAL Patent: EP 0430645-A 1 05-JUN-1991;
 THE WELLCOME FOUNDATION LIMITED
 FEATURES
 SOURCE 1. 1359
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 526 a 149 c 226 g 458 t
 ORIGIN
 Query Match 96.7%; Score 1346.4; DB 6; Length 1359;
 Best Local Similarity 99.6%; Pred. No. 9e-169;
 Matches 1350; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

|||||||
 Db 424 AAGATTCGGGGAGAGATGAGCAAAATTAATCTTTAGGATTAACCGATTAATTTAAT 483
 |||||||
 QY 517 CCTATTATGCAATTAATGCGTTTATTTATCTATTTACTAATGATAGTATTTCTGCT 576
 |||||||
 Db 484 GCTTATTTAGCAATTAATGCGTTTATTTATCTAATTTACTAATGATAGTATTTCTGCT 543
 |||||||
 QY 577 AATTTGATTAATTAATGCGTCTTATGGAAGTGCAGAAATTTACTGTTTGAAGCTATT 636
 |||||||
 Db 544 AATTTGATTAATTAATGCGTCTTATGGAAGTGCAGAAATTTACTGTTTGAAGCTATT 603
 |||||||
 QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 696
 |||||||
 Db 604 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 663
 |||||||
 QY 697 TCTATGATTAATTTAGATATTTTGGCAACGTTAAATCCCAAAAGATGGAATTA 756
 |||||||
 Db 664 TCTATGATTAATTTAGATATTTTGGCAACGTTAAATCCCAAAAGATGGAATTA 723
 |||||||
 QY 757 TACACAGTATTTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
 |||||||
 Db 724 TACACAGTATTTATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
 |||||||
 QY 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
 |||||||
 Db 784 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
 |||||||
 QY 877 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 936
 |||||||
 Db 844 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 903
 |||||||
 QY 937 TATATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
 |||||||
 Db 904 TATATAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
 |||||||
 QY 997 GAAATAGATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1056
 |||||||
 Db 964 GAAATAGATTTTGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
 |||||||
 QY 1057 AATGAGCAATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
 |||||||
 Db 1024 AATGAGCAATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
 |||||||
 QY 1117 CTAAGAGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1176
 |||||||
 Db 1084 CTAAGAGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1143
 |||||||
 QY 1177 TTGCGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1236
 |||||||
 Db 1144 TTGCGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1203
 |||||||
 QY 1237 TTAGAGCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1296
 |||||||
 Db 1204 TTAGAGCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1263
 |||||||
 QY 1297 ATTGCAACCACTGCTACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
 |||||||
 Db 1264 ATTGCAACCACTGCTACTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1323
 |||||||
 QY 1357 TTTGATCTACAGATGAGGATGACCAATGATTA 1392
 |||||||
 Db 1324 TTTGATCTACAGATGAGGATGACCAATGATTA 1359
 |||||||
 RESULT 8
 I28430 1359 bp DNA linear PAT 06-FEB-1997
 LOCUS I28430
 DEFINITION Sequence 1 from patent US 5571694.
 ACCESSION I28430
 VERSION I28430.1 GI:1819206
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

```

Unclassified.
REFERENCE 1 (bases 1 to 1359)
AUTHORS Makoff,A.J., Romanos,M.A., Clare,J.J. and Fairweather,N.F.
TITLE Expression of tetanus toxin fragment C in yeast
JOURNAL Patent: US 5571694-A 1 05-NOV-1996;
FEATURES
    source      1. 1359
                Location/Qualifiers
BASE COUNT    526 a      149 c      226 g      458 t
ORIGIN
Query Match      96.7%; Score 1346.4; DB 6; Length 1359;
Best Local Similarity 99.6%; Pred. No. 9e-169;
Matches 1350; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 37 AAAAATCTGGATTGTTGGTGTGATTAATGAAGACATATAGATTTTAAATGAAAAAGAGT 96
DB 4 AAAAATCTGGATTGTTGGTGTGATTAATGAAGACATATAGATTTTAAATGAAAAAGAGT 63
QY 97 ACAATTTTAAATTTAGATTAATTAATGATTAATATACAGATATATCTGGTTTAATTTCA 156
DB 64 ACAATTTTAAATTTAGATTAATTAATGATTAATTAATGATTAATATCTGGTTTAATTTCA 123
QY 157 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGGAATTAATGGCAAGCAATACAT 216
DB 124 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGGAATTAATGGCAAGCAATACAT 183
QY 217 TTAGTAAACAAATGAACTCTCTGAAGTTAATGTCATTAAGCTATGATTAATTAAT 276
DB 184 TTAGTAAACAAATGAACTCTCTGAAGTTAATGTCATTAAGCTATGATTAATTAAT 243
QY 277 GATATGTTAAATTAATTTTACCTGAGTATTTGGTGGAGGTTCTTAAAGTATCTGCTAGT 336
DB 244 GATATGTTAAATTAATTTTACCTGAGTATTTGGTGGAGGTTCTTAAAGTATCTGCTAGT 303
QY 337 CATTTAGAACAAATATGCGACAAATGAGTATTCATTAATTTAGCTCTATGAAAAAATAGT 396
DB 304 CATTTAGAACAAATATGCGACAAATGAGTATTCATTAATTTAGCTCTATGAAAAAATAGT 363
QY 397 CATATCAATAGGATCTGGTTGGAGTGTATCACTTAAAGTAAATTAATTAATGAGCTTTA 456
DB 364 CATATCAATAGGATCTGGTTGGAGTGTATCACTTAAAGTAAATTAATTAATGAGCTTTA 423
QY 457 AAAGATTCGCGGAGGAGATTAAGCAAAATTAACCTTTAGGATTTTACCTGATTAATTAAT 516
DB 424 AAAGATTCGCGGAGGAGATTAAGCAAAATTAACCTTTAGGATTTTACCTGATTAATTAAT 483
QY 517 GCTTATTTAGCAATTAATGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAAT 576
DB 484 GCTTATTTAGCAATTAATGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAAT 543
QY 577 AATTTGTATTAATTAATGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 636
DB 544 AATTTGTATTAATTAATGAGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 603
QY 637 AGAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
DB 604 AGAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
QY 697 TCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
DB 664 TCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
QY 757 TACACAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
DB 724 TACACAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
QY 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
DB 784 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
QY 877 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936

```

```

DB 844 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
QY 937 TATTTAGAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
DB 904 TATTTAGAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
QY 997 GAAATGATTTCTTTTGTAAATCAGGATTTATTAATTAATTAATTAATTAATTAATTAAT 1056
DB 964 GAAATGATTTCTTTTGTAAATCAGGATTTATTAATTAATTAATTAATTAATTAATTAAT 1023
QY 1057 AATGACCAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
DB 1024 AATGACCAATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1083
QY 1117 CTAGAGTAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1176
DB 1084 CTAGAGTAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
QY 1177 TTGCGTATTTAAAAACCTATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1236
DB 1144 TTGCGTATTTAAAAACCTATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1203
QY 1237 TTAGAGCTATGTTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1296
DB 1204 TTAGAGCTATGTTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
QY 1297 ATTGCAAGCAATGTTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
DB 1264 ATTGCAAGCAATGTTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
QY 1357 TTTGTACCTACAGATGAGGATGAGCAAAATGATTAAT 1392
DB 1324 TTTGTACCTACAGATGAGGATGAGCAAAATGATTAAT 1359

```

```

RESULT 9
AF154828
LOCUS AF154828 1356 bp DNA linear BCT 01-JUN-2000
DEFINITION Clostridium tetani tetanus toxin gene, complete cds.
ACCESSION AF154828
VERSION AF154828.1 GI:8132374
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS He H.J., Shi H.J., He Z.Y., Yuan Q.S. and Wu X.F.
TITLE Fragment C of tetanus toxin
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1356)
AUTHORS He H.J., Shi H.J., He Z.Y., Yuan Q.S. and Wu X.F.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1999) Department of Biochemistry, East China Univ. of Science and Technology, 130 Meilong Road, Shanghai, Shanghai 200237, P.R.China
FEATURES
    source      1. 1356
                Location/Qualifiers

```

```

CDS
    /organism="Clostridium tetani"
    /strain="20886"
    /db_xref="taxon:1513"
    <1..1356
    /note="fragment C"
    /codon_start=1
    /transl_table=11
    /product="tetanus toxin"
    /protein_id="AAE3267.1"
    /db_xref="GI:8132375"
    /translation="KNDQWVDEEDIDVILKSTILINDINDIISDGFNSVIT
    YPDQOLVPGINGKAHLVNNSSPVIVHKAMDIDYNDPNNFTVSPMLRVKVSASHL
    EYDINDVEYSLISGMSKYSLSIGSGMSKYSLSIGSGMSKYSLSIGSGMSKYSLSIG
    AYLANWVFTITNDRLSSANLVINGVMSAEITTGATREDNNITLKDRCNNNGO
    YSIDKRFICKALNPREIEKLYTSLSTIFLRFDEWGNPLRYDTEYVLIIPVASSKDY

```


Db 244 GATATGTTAAATAATTTAACCGTTAGCTTTGGTTGAGGGTTCCTAAAGTATCTGCTAGT 303
 QY 337 CATTAGAACAAATATGGCAACAATGAGTATTCATATATTAGTCTCTATGAAAAACATAGT 396
 Db 304 CATTAGAACAAATATGGCAACAATGAGTATTCATATATTAGTCTCTATGAAAAACATAGT 363
 QY 397 CATTAGAACAAATATGGTGTGGAGTGTATCTATTAAGGTAAATACCTTAATATGAGCTTTA 456
 Db 364 CATTAGAACAAATATGGTGTGGAGTGTATCTATTAAGGTAAATACCTTAATATGAGCTTTA 423
 QY 457 AAAGATTCGGCGGAGAACTTATGACAAATACCTTTAGGATTTAGCTGATTAATTTAAT 516
 Db 424 AAAGATTCGGCGGAGAACTTATGACAAATACCTTTAGGATTTAGCTGATTAATTTAAT 483
 QY 517 GCTTATTTAGCAAAATTAATGGGTTTTTATACCTATTTACTATGATGATTAATGCTTCTGCT 576
 Db 484 GCTTATTTAGCAAAATTAATGGGTTTTTATACCTATTTACTATGATGATTAATGCTTCTGCT 543
 QY 577 AATTTGATTAATTAATGAGTACTTATTTAGGAACTGCGAATTTAGCTGTTAGAGCTAAT 636
 Db 544 AATTTGATTAATTAATGAGTACTTATTTAGGAACTGCGAATTTAGCTGTTAGAGCTAAT 603
 QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
 Db 604 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
 QY 697 TCTATGTTAAATTTAGGATTTTGGCAAGCATTAATTAATTAATTAATTAATTAATTAAT 756
 Db 664 TCTATGTTAAATTTAGGATTTTGGCAAGCATTAATTAATTAATTAATTAATTAATTAAT 723
 QY 757 TACCAAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
 Db 724 TACCAAGTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
 QY 817 GATACAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
 Db 784 GATACAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
 QY 877 ATACAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
 Db 844 ATACAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903
 QY 937 TATATGAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
 Db 904 TATATGAGAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
 QY 997 GAAATAGATTCCTTTGTTAAATCAGGTGATTTTATTAATTAATTAATTAATTAATTAAT 1056
 Db 964 GAAATAGATTCCTTTGTTAAATCAGGTGATTTTATTAATTAATTAATTAATTAATTAAT 1023
 QY 1057 AATGAGCAATTTAGGATTTTCCGAAAGATGGAATGCTTTAATTAATTAATTAATTAATTAAT 1116
 Db 1024 AATGAGCAATTTAGGATTTTCCGAAAGATGGAATGCTTTAATTAATTAATTAATTAATTAAT 1083
 QY 1117 CTAAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1176
 Db 1084 CTAAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1143
 QY 1177 TTGCGTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1236
 Db 1144 TTGCGTATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1203
 QY 1237 TTGAGGATTTAGGATTTTCCGAAAGATGGAATGCTTTAATTAATTAATTAATTAATTAATTAAT 1296
 Db 1204 TTGAGGATTTAGGATTTTCCGAAAGATGGAATGCTTTAATTAATTAATTAATTAATTAATTAAT 1263
 QY 1297 ATTGCAAGCAATGCTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
 Db 1264 ATTGCAAGCAATGCTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
 QY 1357 TTTTGTACCTACAGATGAGGATGAGCAAAATGATTTAA 1392
 Db 1324 TTTTGTACCAACGATGAGGATGAGCAAAATGATTTAA 1359

RESULT 11
 A18201 1359 bp DNA linear PART 09-JUL-2002
 LOCUS synthesised fragment C DNA increased (G+C) content seq ID No:2.
 DEFINITION A18201
 ACCESSION A18201
 VERSION A18201.1 GI:212727091
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 artificial sequences.
 1 (bases 1 to 1359)
 AUTHORS
 TITLE
 JOURNAL
 Makoff,A.J., Romanos,M.A., Clare,J.J. and Fairweather,N.F.
 Patent: EP 0430645-A 3 05-JUN-1991;
 THE WELLCOME FOUNDATION LIMITED
 FEATURES
 source
 1..1359
 location/Qualifiers
 BASE COUNT 381 a 374 c 271 g 333 t
 ORIGIN
 Query Match 57.4%; Score 799.2; DB 6; Length 1359;
 Best Local Similarity 74.3%; Pred. No. 1.2e-96;
 Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;
 QY 37 AAAAATCTGATTTGGTGGTGAATTAATGAAGATATAGATGTTATTAATTAATTAATTAATTAAT 96
 Db 4 AAAAATCTGATTTGGTGGTGAATTAATGAAGATATAGATGTTATTAATTAATTAATTAATTAAT 63
 QY 97 ACAATTTAAATTTAGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 156
 Db 64 ACAATTTAACTTGAATTCGATTAACAAACAGATATTAATTAATTAATTAATTAATTAATTAAT 123
 QY 157 TCTGTATTAACATATCCAGATGCTCAATTTGGTGGCGGAAATTAATTAATTAATTAATTAAT 216
 Db 124 TCTGTATTAACATATCCAGATGCTCAATTTGGTGGCGGAAATTAATTAATTAATTAATTAAT 183
 QY 217 TTTGTAACATGATTAATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 276
 Db 184 TTTGTAACATGATTAATTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 243
 QY 277 GATATGTTTAAATTTATTTAGCTTTAGCTTTGTTGAGGCTTCTTAAGTATTTCTGCTAGT 336
 Db 244 GATATGTTTAACTTACCTGTTAGCTTTGTTGAGGCTTCTTAAGTATTTCTGCTAGT 303
 QY 337 CATTAGAACAAATATGGCAACAATGAGTATTCATATATTAGTCTCTATGAAAAACATAGT 396
 Db 304 CATTAGAACAAATATGGCAACAATGAGTATTCATATATTAGTCTCTATGAAAAACATAGT 363
 QY 397 CATTAGAACAAATATGGTGTGGAGTGTATCTATTAAGGTAAATACCTTAATATGAGCTTTA 456
 Db 364 CATTAGAACAAATATGGTGTGGAGTGTATCTATTAAGGTAAATACCTTAATATGAGCTTTA 423
 QY 457 AAAGATTCGGCGGAGAACTTATGACAAATACCTTTAGGATTTAGCTGATTAATTTAAT 516
 Db 424 AAAGATTCGGCGGAGAACTTATGACAAATACCTTTAGGATTTAGCTGATTAATTTAAT 483
 QY 517 GCTTATTTAGCAAAATTAATGGGTTTTTATACCTATTTACTATGATGATTAATGCTTCTGCT 576
 Db 484 GCTTATTTAGCAAAATTAATGGGTTTTTATACCTATTTACTATGATGATTAATGCTTCTGCT 543
 QY 577 AATTTGATTAATTAATGAGTACTTATTTAGGAACTGCGAATTTAGCTGTTAGAGCTAAT 636
 Db 544 AATTTGATTAATTAATGAGTACTTATTTAGGAACTGCGAATTTAGCTGTTAGAGCTAAT 603
 QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
 Db 604 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
 QY 697 TCTATGTTAAATTTAGGATTTTGGCAAGCATTAATTAATTAATTAATTAATTAATTAAT 756

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
BASE COUNT
ORIGIN

unidentified.
unclassified.
1 (bases 1 to 4366)
Khan, M.A., Hormeache, C.E., Chatfield, S.N. and Dougan, G.
VACCINE COMPOSITIONS EXPRESSION OF RECOMBINANT FUSION PROTEINS IN
ATTENUATED BACTERIA
Patent: WO 9504151-A 12 09-FEB-1995.
MEDIVA HOLDINGS B V (NL)
Other publication PL 313979 960805
Other publication GB 2295394 960529
Other publication JP 8503602P 960423
Other publication CA 2168459 950209
Other publication NO 950348 950328
Other publication AU 7235794 950228
Other publication FI 950396 950130
Other publication AU 4719393 940303.
Location/Qualifiers
1..4366
/organism="unidentified"
/db.xref="taxon:32644"
1176 a 1081 c 1011 g 1098 t

Query Match 57.3%; Score 797.2; DB 6; Length 4366;
Best Local Similarity 74.3%; Pred. No. 1.5e-96;
Matches 1006; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

37 AAAAAATCGATTTGGTGGTGATATATGACAGATATAGATGTTATATTAATAAGAGT 96
101 AAAAACTTGTATGTTGGTGCAGACAGAGAGACATGATGTTATCTGTAAGAAAGTCT 160
97 ACAATTTTAAATTTGATATTAATATATATATATATATATATATATATATATATATAT 156
161 ACCATTTCTGMACTTGACATCAACAGACATATATATCTCGACATCTCGTTTCACTCC 220
157 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGCAATTAATGCAAGCAATACAT 216
221 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGCAATTAATGCAAGCAATACAT 280
217 TTAGTAACATGATCTTCTGCAATTTAGTCAATTAAGCATGATGATATTAATATATAT 276
281 CTGCTTAACAGATCTTCTGCAATTTAGTCAATTAAGCATGATGATATTAATATATAT 340
277 GATATGTTAATATATTTACGTTAGCTTTGGTGAAGGTTCTTAAGTATCTGCTAGT 336
341 GACATGTTCAACACTTCACTGCTTGTAGCTTGTGCGGCTCCGAAAGTTTCTGCTCC 400
337 CATTTAGAACATATGACCAATAGATTCATTAATTAAGTCTATGAAAAACATAGT 396
401 CACCTGGAACATGACGCACTAGACAGTACCTATCATCATGCTCTATGAGAAACACTCC 460
397 CTATCAATAGATCTGTTGGAGTATCACTTAAGAGTAATTAATTAATTAATGAATTA 456
461 CTGCTCAATGCTGCTGTTGGTCTGTTCCCTGAAGGGTAACACCTGATCTGACTCG 520
457 AAGATTTCCGCGGAGAGTGAACAATTAATTTAGGATTTACCTGATTAATTAAT 516
521 AAGATTTCCGCGGAGAGTGAACAATTAATTTAGGATTTACCTGATTAATTAAT 580
517 GCTTATTTAGCAATTAATGAGTTTATTAATTAATTAATTAATTAATTAATTAATTA 576
581 GGTATCTGCTGCTAACAATGAGTTTATTAATTAATTAATTAATTAATTAATTAATTA 640
577 AATTTGATATTAATGAGTACTTATGGAAGTGAACAATTAATTAATTAATTAATTAAT 636
641 AACCTGATATCAACGAGCTTCTGATGGCTCCGCTGAATCACTGCTGCTGCTGCTATC 700
637 AAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
701 CGTGAAGACACATCACTCTTAAGCTGGAACGTTGCAACACACACACACACACATGTA 760

697 TCTATGATTAATTTAGGATATTTTGCAGACATTAATCCAAAAGAGATTGAAAAATTA 756
761 TCCATTCGCAAGTTCCGATATCTTCGCAAGACACTGAAACCCGAAAGAGTGAACAACTG 820
757 TACACAGATTTATATATACCTTTTAAAGACATCTGGGGAACCCCTTTGAGATAT 816
821 TATACACACTTCCGCTATATACCTTCTGCTGATCACTTGGGTAACCCGCTGCTTAC 880
817 GATACAGATTTATTTATATATACAGTCTTCTGATTTAAAGATTTGTAATTTGAAAAAT 876
881 GACACCGAATTTACGATATCCCGGATCTTACGCTTAAGACCTTACAGCTGAAAAAC 940
877 ATACAGATTTATGATTTTAAACAATGCGCATCTTACTAACGGAATTTGATATA 936
941 ATCACTGATCAATGATCTGATCAACACGCGCTTCACTTAAGCTGAACCTGAACATC 1000
937 TATATACAGATTTATATATGAGTAAATTTATTAATAAGATTTACACCTTAATAT 996
1001 TACTACCGACGCTGTATACAGGCTGAAATTCATATCAATCAACGCTACACCTCCGAAC 1060
997 GAATAGATTTCTTTGTTAAATCAGTATTTTATTAATTAATTAATTAATTAATTAAT 1056
1061 GAATAGATTTCTTTGTTAAATCAGTATTTTATTAATTAATTAATTAATTAATTAAT 1120
1057 AATGACGATTTGATGATTTATCCGAAGATGAATGCTTTAATTAATCTGATGAAT 1116
1121 AAGCAACATCTGTTGTTACCGAAAGCGTAAACGCTTCAACACCTGACAGAAAT 1180
1117 CTAAAGATGATTTAATGAGTATCCCTCTTCTTAAATAAGATGAAGCAGTAA 1176
1181 CTGCTGTTGTTGTTACAGGCTCCGGTATCCCGCTACAAAAATGGAAGCTTTAA 1240
1177 TTGCTGATTTAAACCTATTTCTTACACTTAATTAATTAATTAATTAATTAATTAAT 1236
1241 CTGCTGATCTGAAACCTTACCTGCTGATGAAACGTTACAGACGAAACAAACCTTCT 1300
1237 TTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1296
1301 CTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1360
1297 AATGCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1356
1361 ATGCTTCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1420
1357 TTTGATCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390
1421 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1454

RESULT 15
A37074
LOCUS A37074 3754 bp DNA circular PAT 05-MAR-1997
DEFINITION Sequence 17 from Patent WO9403615.
ACCESSION A37074
VERSION A37074.1 GI:2294248
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

unidentified.
unclassified.
1 (bases 1 to 3754)
Dougan, G.
Khan, M.A., Hormeache, C.E., Villarreal-Ramos, B., Chatfield, S.N. and
EXPRESSION OF RECOMBINANT FUSION PROTEINS IN ATTENUATED BACTERIA
Patent: WO 9403615-A 17 17-FEB-1994;
MEDIVA HOLDINGS B V (NL)
Other publication CA 2141477 940217
Other publication AU 719393 940303
Other publication NO 950348 950328
Other publication FI 950396 950130
Other publication JP 8503602P 960423.
Location/Qualifiers
1..3754
/organism="unidentified"

BASE COUNT	978 a	977 c	860 g	939 t
ORIGIN	/db_xref="taxon:32644"			

Query Match	57.28;	Score 796.8;	DB 6;	Length 3754;
Best Local Similarity	74.38;	Pred. No. 1.8e-96;		
Matches 1005; Conservative	0;	Mismatches 347;	Indels	0.

[illegible]

Db	1061	GAATGCAATCTTTCGTTAAATCTGGGAGACTTCATCAAACTGTCGTTCTTACACAAAC	1120
Qy	1057	AATGAGCACATTTAGGTTATCCGAAAGATGGAATACTTATATCTTGATGAAT	1116
Db	1121	AAGAACACATCGTTGGTTACCGAAAGACGGTAAAGCTTTCAACAACCTGGACGAAT	1180
Qy	1117	CTAAGAGTAGTATATATAGCCGACGAGTATCCCTTTATAAAAAATGAAGCAGTAAA	1176
Db	1181	CTGCGGTGTGGTTACAAAGCTCCGGGATCCGCTGTACAAAAAATGGAAGCTGTAAA	1240
Qy	1177	TTCGCTGATTTAAAAACCTATTCCTGTACAACTTAAATATATGATGATTAATAATGCATCT	1236
Db	1241	CTGCGAGACCTGAAAAACCTACTCTGTTACCTGAACCTGTACGACGACAAAAACGCTTCT	1300
Qy	1237	TTAGAGCTAGTAGTACGATCCCATATATGTCAAATATAGCACAGATCCAAATAGGATATATATTA	1296
Db	1301	CTGGGTCTGGTTGTTACCCACAACGGTCAAGATCGTATACGACCCGACCGTGCATCTCTG	1360
Qy	1297	ATTGCACGACACTGGTACTTTAATCATTTTAAAAAGATAAAAATTTAGAGTGTGATTTGGTAC	1356
Db	1361	ATTCGCTTCTAACTGTACTTCAACCAACCACTGAAAGACAAAATCTCGGTTGCCAGCTGTAC	1420
Qy	1357	TTTGTACTACGATGAGGAGATGAGCAAAATGA	1388
Db	1421	TTTCGTTCCGACCGATGAAGGTTGGACCAACGA	1452

```
Search completed: December 1, 2002, 09:49:39
Job time : 2749 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:53:27 ; Search time 254 Seconds

(Without alignments)
12341.666 Million cell updates/sec

Title: US-09-816-467-3

Sequence: 1 atggtttttcaaccacat.....aggatggacaatgattaa 1392

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1392	100.0	1600	20	AAx27234	Tetanus toxin frag
2	1387.4	99.7	1851	8	AAAT0545	Sequence encoding
3	1379.8	99.1	1858	19	AAV32580	DNA encoding a SOD
4	799.2	57.4	1359	12	AAQ12121	Synthetic tetanus
5	799.2	57.4	3712	16	AAQ97490	Plasmid pTRAI. N
6	796.8	57.3	4366	16	AAQ85424	Plasmid pTRCH3-p28
7	796.8	57.2	3754	15	AAQ57880	Intermediate plasm
8	796.8	57.2	3754	16	AAQ85420	Plasmid pTRCH1. S
9	796.8	57.2	3769	15	AAQ57881	Intermediate plasm

10	796.8	57.2	3769	16	AAQ97492	Plasmid pTRCH2. N
11	796.8	57.2	4377	16	AAQ85421	Plasmid pTRCH1-p28
12	521	37.4	1766	16	AAQ97491	Plasmid pTRCH2 + h
13	248.2	17.8	1535	19	AAV30596	Clostridium botuli
14	234.8	16.9	1547	19	AAV30581	Clostridium botuli
15	233.2	16.8	3876	19	AAV30579	Clostridium botuli
16	232.7	16.7	1526	19	AAV30580	Clostridium botuli
17	228.2	16.4	1317	21	AA287220	DNA encoding nativ
18	228.2	16.4	1546	19	AAV30575	Clostridium botuli
19	228.2	16.4	2532	17	AA287218	DNA encoding nativ
20	228.2	16.4	3891	17	AA287218	DNA encoding nativ
21	228.2	16.4	4835	21	AA287218	DNA encoding nativ
22	220.8	15.9	3509	19	AAV26289	C. botulinum type
23	210.2	15.1	1472	19	AAV30585	BONT/A neurotoxin
24	206.2	14.8	1293	18	AAV30585	Recombinant botuli
25	205.2	14.7	4017	22	AAV48160	Clostridium botuli
26	205	14.7	1463	19	AAV30584	Immunogenic type F
27	176.2	12.7	1460	19	AAV30584	C. botulinum BONT/E
28	163.6	11.8	1917	21	AAV30593	Clostridium botuli
29	152.6	11.0	1469	19	AA287219	DNA encoding nativ
30	113.8	8.2	3950	20	AAV30591	DNA encoding nativ
31	109	7.8	1502	20	AAV30591	Clostridium botuli
32	93.8	6.7	1278	22	AAV30588	DNA coding for mod
33	93.8	6.7	1400	21	AA287215	Clostridium botuli
34	93.8	6.7	1400	22	AAV54489	Botulinum toxin hea
35	85	6.1	1830	24	ABL56243	DNA encoding synth
36	85	6.1	3292	24	ABL56203	Botulinum toxin hea
37	84.2	6.0	15548	24	ABL34155	AmEPV tk DNA. Ams
38	75.6	5.4	1511	13	AAQ28302	AmEPV genome fragm
39	75.6	5.4	1511	15	AAQ66798	AmEPV tk DNA. Ams
40	75.6	5.4	1511	19	AAV14508	AmEPV genome fragm
41	75.6	5.4	1511	20	AA210082	AmEPV genome fragm
42	75.6	5.4	15548	24	ABL34155	AmEPV genome fragm
43	75.6	5.4	5000	24	ABL35643	Human immune syste
44	74.8	5.4	6175	24	ABL35307	Human immune syste
45	74.6	5.4	50000	24	ABL56201	AmEPV genome fragm

ALIGNMENTS

RESULT 1						
ID	AAx27234	standard; DNA; 1600 BP.				
XX	AAx27234:					
AC	XX					
XX	XX					
DF	28-MAY-1999	(first entry)				
XX	XX					
DE	Tetanus toxin fragment C coding sequence.					
KW	Tetanus toxin fragment C; TTC; central nervous system; CNS; spinal cord; proteolytic fragment; retrograde axonal transport; spinal cord disease;					
KW	transynaptic transport; neurodegenerative disease; motoneuron disease;					
KW	amyotrophic lateral sclerosis; spinal muscular atrophy; therapy; ALS;					
XX	SMA; neurodegenerative lysosomal storage disease; neuronal maping; ss.					
OS	Clostridium tetani.					
XX	XX					
PN	WO9909057-A2.					
PD	XX					
XX	25-FEB-1999.					
PF	12-AUG-1998;	98WO-EP05113.				
PR	13-NOV-1997;	97US-0065236.				
PR	14-AUG-1997;	97US-0055615.				
XX	XX					
PA	(INSP) INST PASTEUR.					
XX	XX					
PI	Brulet P, Coen L, Osta Pinzolas R;					
XX	XX					
DR	WPI; 1999-180971/15.					

DR P-PSDB: AAY00921.

XX Delivery of a composition to the central nervous system or spinal
PT cord - comprises administration of a non-toxic, proteolytic
PT fragment of tetanus toxin in association with a molecule having
PT biological function

PS Example 1: Fig 1: 53bp; English.

XX This sequence encodes the tetanus toxin fragment C (TTC):
CC The invention relates to a method for in vivo delivery of a desired
CC composition into a human or animal central nervous system (CNS) or spinal
CC cord comprising administering a non-toxic, proteolytic fragment of tetanus
CC toxin (TTC) in association with at least a molecule having a biological
CC function and where the composition is capable of in vivo retrograde
CC axonal transport and transsynaptic transport into the CNS or the spinal
CC cord of the human or animal and of being delivered to different areas of
CC the CNS or the spinal cord. The method can be used for the treatment of
CC humans or animals with CNS or spinal cord disease, e.g. neurodegenerative
CC and motoneuron diseases such as amyotrophic lateral sclerosis (ALS),
CC spinal muscular atrophies (SMA) or neurodegenerative lysosomal storage
CC diseases. Compositions comprising hybrid fragments of TTC comprising
CC fragments C and B can also be used for neuronal mapping and
CC immunisations. Use of TTC comprising fragments A, B and C results in
CC better transport of the fragment inside the organism compared with
CC fragment C.

XX Sequence 1600 BP; 589 A; 213 C; 278 G; 520 T; 0 other;

Query Match 100.0%; Score 1392; DB 20; Length 1600;
Best Local Similarity 100.0%; Pred. No. 4.3e-230;

Matches 1392; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTTTTTCACACCAATTCATTTCTTAAATTCGATGTTGGGTTGAT 60
DB 88 ATGCTTTTTCACACCAATTCATTTCTTAAATTCGATGTTGGGTTGAT 147
QY 61 AATGAAGAAGATATAGATATATTAATAAGATCAATTTAAATTTGATATTAAT 120
DB 148 AATGAAGAAGATATAGATATATTAATAAGATCAATTTAAATTTGATATTAAT 207
QY 121 AATGATATATATAGATATATCTGGTTTAAATCATCTGTAATACATATCCAGATCT 180
DB 208 AATGATATATATAGATATATCTGGTTTAAATCATCTGTAATACATATCCAGATCT 267
QY 181 CAATTGGTCCCGGAATTAATGGCAACCAATACATTTAGTAACCAATGATCTTGAA 240
DB 268 CAATTGGTCCCGGAATTAATGGCAACCAATACATTTAGTAACCAATGATCTTGAA 327
QY 241 GTTATAGTCATTAAGCTATGATATTAATTAATTAATTAATTTACCGTT 300
DB 328 GTTATAGTCATTAAGCTATGATATTAATTAATTAATTTACCGTT 387
QY 301 AGCTTTTGGTGGGTTCTTAAGATCTGCTAGTCAATTTGAACAATATGGCAATAT 360
DB 388 AGCTTTTGGTGGGTTCTTAAGATCTGCTAGTCAATTTGAACAATATGGCAATAT 447
QY 361 GAGTATTCATTAATTAAGCTATGATATTAATTAATTAATTTACCGTT 420
DB 448 GAGTATTCATTAATTAAGCTATGATATTAATTAATTTACCGTT 507
QY 421 GATACCTTAAGGATTAATTAATTAATTAATTAATTTACCGTT 480
DB 508 GATACCTTAAGGATTAATTAATTAATTAATTTACCGTT 567
QY 481 CAATTAATTAATTAAGGATTAATTAATTAATTTACCGTT 540
DB 568 CAATTAATTAATTAAGGATTAATTAATTTACCGTT 627
QY 541 TTTTATTAATTAATTAAGGATTAATTAATTTACCGTT 600
DB 628 TTTTATTAATTAATTAAGGATTAATTTACCGTT 687

QY 601 ATGGGAAGTGCAGAAATTAATCTGTTAGAGCTATTAAGAGAGATTAATTAATTAACATTA 660
DB 688 ATGGGAAGTGCAGAAATTAATCTGTTAGAGCTATTAAGAGAGATTAATTAATTAACATTA 747
QY 661 AAACATAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 720
DB 748 AAACATAGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 807
QY 721 TGCAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 780
DB 808 TGCAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 867
QY 781 TTTTAAAGATCTTGGGGAACCCCTTACGATATGATACAGAAATTAATTAATTAATTAATTA 840
DB 868 TTTTAAAGATCTTGGGGAACCCCTTACGATATGATACAGAAATTAATTAATTAATTAATTA 927
QY 841 GTAGCTTCTAGTCTTAAGATGTTCAATTTGAAAATTAATTAATTAATTAATTAATTAATTA 900
DB 928 GTAGCTTCTAGTCTTAAGATGTTCAATTTGAAAATTAATTAATTAATTAATTAATTAATTA 987
QY 901 AATGGCCATCGTATACGGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 960
DB 988 AATGGCCATCGTATACGGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1047
QY 961 CTAAATTTTATTAATAAAGATATACACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1020
DB 1048 CTAAATTTTATTAATAAAGATATACACCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1107
QY 1021 GGTATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1080
DB 1108 GGTATATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1167
QY 1081 AAGATGGAATTCCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
DB 1168 AAGATGGAATTCCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1227
QY 1141 GGTATCCCTCTTTTAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1200
DB 1228 GGTATCCCTCTTTTAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1287
QY 1201 GTACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1260
DB 1288 GTACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1347
QY 1261 GGTCAATTAAGCAACGATCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1320
DB 1348 GGTCAATTAAGCAACGATCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1407
QY 1321 CATTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1380
DB 1408 CATTTAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1467
QY 1381 ACAAAATGATTA 1392
DB 1468 ACAAAATGATTA 1479

RESULT 2
AAN70545
ID AAN70545 standard; DNA; 1851 BP.
AC AAN70545;
XX 22-APR-1991 (first entry)
DT Sequence encoding a portion of the B fragment and all of the C
DE fragment of tetanus toxin.
XX TT; vaccine; ds.
XX Clostridium tetani.
OS
XX
FH Key Location/Qualifiers


```

FT CDS 1..1722
XX /*tag= a
XX
XX EP209281-A.
XX
XX 21-JAN-1987.
XX
XX 27-JUN-1986; 86EP-0305029.
XX
XX 28-JUN-1985; 85GB-0016442.
XX
XX (WELL ) WELLCOME FOUNDATION LTD.
XX
XX Fairweather NF:
XX
XX WPI; 1987-015999/03.
XX
XX P-PSDB; AAP70345.
XX
XX Cloned DNA sequence coding for tetanus toxin - or its fragments
XX contg. epitope used to express antigens for vaccine production.
XX
XX Claim 4; Fig 1; 36pp; English.
XX
XX Gene product comprises a tetanus toxin fragment, which may be
XX expressed in a transformed host, and used as an antigen in vaccine
XX production, against the disease.
XX
XX Sequence 1851 BP; 754 A; 204 C; 279 G; 614 T; 0 other:
XX
Query Match 99.7%; Score 1387.4; DB 8; Length 1851;
Best Local Similarity 99.9%; Pred. No. 2.7e-229;
Matches 1388; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 604 GGAAGTGCAGAAATTTACTGGTTAGAGCTATTAGAGGSAATTAATATAACATTTAAA 663
DB 934 GGAAGTGCAGAAATTTACTGGTTAGAGCTATTAGAGGSAATTAATATAACATTTAAA 993
OY 664 CTAGATAGATGTAATTAATTAATCAATAGCTTCTATTGATTAATTTAGGATTTTGC 723
DB 994 CTAGATAGATGTAATTAATTAATCAATAGCTTCTATTGATTAATTTAGGATTTTGC 1053
OY 724 AAAGCATTTAAATCCAAAGAGATTGAAAATTTATACACAGTATTATTATACCTTT 783
DB 1054 AAAGCATTTAAATCCAAAGAGATTGAAAATTTATACACAGTATTATTATACCTTT 1113
OY 784 TTAGAGACTTCTGGGAAACCCCTTACGATATGATACAGATATTTATATACCACTA 843
DB 1114 TTAGAGACTTCTGGGAAACCCCTTACGATATGATACAGATATTTATATACCACTA 1173
OY 844 GCTTCTAGTTCCTAAGATGCTCAATTTGAAAATTAACAGATTAATATTTGACAAAT 903
DB 1174 GCTTCTAGTTCCTAAGATGCTCAATTTGAAAATTAACAGATTAATATTTGACAAAT 1233
OY 904 GCGCCATCGTATCTAACGGAATTTGAATTAATATATATAGAGCTTATATAGACTA 963
DB 1234 GCGCCATCGTATCTAACGGAATTTGAATTAATATATAGAGCTTATATAGACTA 1293
OY 964 AAATTTATTTAAAGAAATATACACCTAATATGAATAGATTCTTTTGTAAATCAGT 1023
DB 1294 AAATTTATTTAAAGAAATATATACACCTAATATGAATAGATTCTTTTGTAAATCAGT 1353
OY 1024 GATTTTATTTAAATTTATGATATATATACAAATATAGAGACATTTAGGTTATCCGAA 1083
DB 1354 GATTTTATTTAAATTTATGATATATATACAAATATAGAGACATTTAGGTTATCCGAA 1413
OY 1084 GATGGAATGCTTTATATATCTTGAAGATCTTAAGATAGTATATGCCCCAGT 1143
DB 1414 GATGGAATGCTTTATATATCTTGAAGATCTTAAGATAGTATATGCCCCAGT 1473
OY 1144 ATCCCTCTTTTAAAGAAATTTGAAGCAGTAAATTTGCGATTTAAACCTATTCGTA 1203
DB 1474 ATCCCTCTTTTAAAGAAATTTGAAGCAGTAAATTTGCGATTTAAACCTATTCGTA 1533
OY 1204 CAACCTAAATTTATGATATATATATATATATATATATATATATATATATATATAT 1263
DB 1534 CAACCTAAATTTATGATATATATATATATATATATATATATATATATATATATATAT 1593
OY 1264 CAATATAGCAACGATCAATATAGGATATATATATATATATATATATATATATATATAT 1323
DB 1594 CAATATAGCAACGATCAATATAGGATATATATATATATATATATATATATATATATAT 1653
OY 1324 TTTAAAGATTTAAATTTTATGATGATGTTGTAATTTGTAACCTACAGATGAGGATGACA 1383
DB 1654 TTTAAAGATTTAAATTTTATGATGATGTTGTAATTTGTAACCTACAGATGAGGATGACA 1713
OY 1384 AATGATTTAA 1392
DB 1714 AATGATTTAA 1722

```

RESULT 3
AAV32580
ID AAV32580 standard; DNA; 1658 BP.
XX

AAV32580;
XX
23-SEP-1998 (first entry)
XX

DNA encoding a SOD-1/TTC hybrid protein.
XX

Chimeric; copper-zinc superoxide dismutase; SOD-1; TTC; SOD; Ret451;
KW tetanus toxin fragment C; tetanus holotoxin; nerve cell; stroke;
KW neurological disorder; oxidative stress; brain hypoxia-reperfusion;
KW epilepsy; Parkinson's disease; Huntington's disease; ss.
XX
OS Chimeric - Homo sapiens.

XX	Chimeric - Clostridium tetani.
XX	Key
XX	Location/Qualifiers
XX	2..1858
XX	/tag= a
XX	/product= "SOD-1/TTC hybrid protein"
XX	2..490
XX	/tag= b
XX	/note= "encodes SOD-1 region of the SOD-1/TTC hybrid protein"
XX	503..1855
XX	/tag= c
XX	/note= "encodes TTC region of the SOD-1/TTC hybrid protein"
XX	US5780024-A.
XX	14-JUL-1998.
XX	21-JUN-1996; 96US-0668381.
XX	23-JUN-1995; 95US-0000473.
XX	21-JUN-1996; 96US-0668381.
XX	(GEHO) GEN HOSPITAL CORP.
XX	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	Brown RH, Fishman PS, Francis JW, Hosler BA;
XX	WPI: 1998-412999/35.
XX	P-PSDB: AAM48909.
XX	New hybrid protein of superoxide dismutase and tetanus toxin fragment C - having increased uptake by neurons and retention of enzymatic activity in these cells, for treating neurological diseases associated with oxidative stress
XX	Disclosure: Columns 25-28: 23pp: English.
XX	The present DNA sequence encodes an enzymatically active human copper-zinc superoxide dismutase (SOD-1) fused at its carboxyl terminus with the tetanus toxin fragment C (TTC) moiety. The TTC moiety constitutes amino acid residues 865-1315 of the tetanus holotoxin. The hybrid protein, referred as SOD:ret451, is claimed to have the following properties: (a) it exhibits Cu/Zn SOD enzymatic activity; (b) the TTC moiety selectively binds to nerve cells and allows uptake of the hybrid protein into these cells; and (c) it retains substantial SOD enzymatic activity following cellular uptake. SOD:ret451 is claimed to be useful for treating neurological disorders associated with oxidative stress, e.g. stroke, brain hypoxia-reperfusion, epilepsy, Parkinson's and Huntington's diseases.
XX	Sequence 1858 BP; 665 A; 248 C; 373 G; 572 T; 0 other;
XX	Query Match 99.1%; Score 1379.8; DB 19; Length 1858;
XX	Best Local Similarity 99.9%; Pred. No. 5.3e-228; Indels 0; Gaps 0
XX	Matches 1381; Conservative 0; Mismatches 2;
XX	10 TCACACCAATTCATTTCTATTCTTAAAAATCTGGATTGGTGGTGAATGAAGAA 69
XX	
XX	Best Local Similarity 99.9%; Pred. No. 5.3e-228; Indels 0; Gaps 0
XX	Matches 1381; Conservative 0; Mismatches 2;
XX	476 TCACACCACTTCATTTCTATTCTTAAAAATCTGGATTGGTGGTGAATGAAGAA 535
XX	
XX	70 GATATAGATGTATATTTAAAAAGAGACAAATTTAAATTAAGTAAATGAATAT 129
XX	
XX	536 GATATAGATGTATATTTAAAAAGAGACAAATTTAAATTAAGTAAATGAATAT 595
XX	
XX	130 ATATCAGATATATCTGGGTTAAATTCATCTGTAATACATATCCAGATGCTCAATGGTG 189
XX	
XX	596 ATATCAGATATATCTGGGTTAAATTCATCTGTAATACATATCCAGATGCTCAATGGTG 655
XX	
XX	190 CCGGAATTAATGGCAAGCAATACATTTAGTAAGCAATGAATCTTGGAGAGTATATGTG 249
XX	
XX	656 CCGGAATTAATGGCAAGCAATACATTTAGTAAGCAATGAATCTTGGAGAGTATATGTG 715

Qy	250	CATAAAGCTATGAGATATTGAAATATATGATATAGTTTAATATTTTAACGTTAGCTTTGG	309
Db	716	CATAAAGCTATGAGATATTGAAATATATGATATAGTTTAATATTTTAACGTTAGCTTTGG	775
Qy	310	TTGAGGGTTCCTAAAGTATCTGCTAGTCCTTTAGAAACATATGCGCAAAATGACTTCA	369
Db	776	TTGAGGGTTCCTAAAGTATCTGCTAGTCCTTTAGAAACATATGCGCAAAATGACTTCA	835
Qy	370	ATAAATTAGCTCTATGAAAAAAACATAGTCTATCAATATGATCTGGTTGAGTATCACTT	429
Db	836	ATAATTAGCTCTATGAAAAAAACATAGTCTATCAATATGATCTGGTTGAGTATCACTT	895
Qy	430	AAAGTAACTACTTATATATGACCTTTAAAGATTCGCGGGAGAGTTGACAAATAACT	489
Db	896	AAAGTAAATAACTTATATATGACCTTTAAAGATTCGCGGGAGAGTTGACAAATAACT	955
Qy	490	TTTAGGAGATTACCTGATGAATTTTATAGCTTATTTAGCAAAATAAGGGTTTATTAAC	549
Db	956	TTTAGGAGATTACCTGATGAATTTTATAGCTTATTTAGCAAAATAAGGGTTTATTAAC	1015
Qy	550	ATTACTAATGATAGATTATCTCTGCTAATTTTGTATATTAATGGAGTACTTATGGAACT	609
Db	1016	ATTACTAATGATAGATTATCTCTGCTAATTTTGTATATTAATGGAGTACTTATGGAACT	1075
Qy	610	GCAGAAATTAAGTGGTTTAGAGCTATTACAGAGGATAATATATACATTTAAACATACAT	669
Db	1076	GCAGAAATTAAGTGGTTTAGAGCTATTACAGAGGATAATATATACATTTAAACATACAT	1135
Qy	670	AGATGTAATTAATTAATTAATCAATACGTTTCTATTTGATTAATTAATTAATGATATTTTGC	729
Db	1136	AGATGTAATTAATTAATTAATCAATACGTTTCTATTTGATTAATTAATTAATGATATTTTGC	1195
Qy	730	TTTAATCCAAAAGAGTTGAAAAATTAATACAGAGTTATTTATCTATTAACCTTTTAAG	789
Db	1196	TTTAATCCAAAAGAGTTGAAAAATTAATACAGAGTTATTTATCTATTAACCTTTTAAG	1255
Qy	790	GACTTGTGGGAAACCTTTAGCATATGATACAGAAATATTTAATTAATACAGTACTTCT	849
Db	1256	GACTTGTGGGAAACCTTTAGCATATGATACAGAAATATTTAATTAATACAGTACTTCT	1315
Qy	850	AGTTCTAAAGATCTTCAATTTGAAAAATATATACAGATTATATGATATTGACAAAATGGGCA	909
Db	1316	AGTTCTAAAGATCTTCAATTTGAAAAATATATACAGATTATATGATATTGACAAAATGGGCA	1375
Qy	910	TGCTATACATACGGAATTTGAT	969
Db	1376	TGCTATACATACGGAATTTGAT	1435
Qy	970	ATTATATAAAGATATACACTTAATTAATGAATAGATTCTTTTGTAAATAGAGTATTT	1029
Db	1436	ATTATATAAAGATATACACTTAATTAATGAATAGATTCTTTTGTAAATAGAGTATTT	1495
Qy	1030	ATTAAATTTATATGATATATATTAACATTAATGACACACTTGTAGTATTCGCAAAAGATGA	1089
Db	1496	ATTAAATTTATATGATATATTAACATTAATGACACACTTGTAGTATTCGCAAAAGATGA	1555
Qy	1090	AATGCTTTTAATATCTTTGATAGAAATTTCTAAGAGTAGATTATATAGTCCCAAGTATTCCT	1149
Db	1556	AATGCTTTTAATATCTTTGATAGAAATTTCTAAGAGTAGATTATATAGTCCCAAGTATTCCT	1615
Qy	1150	CTTTATTAATAAATTTGGAAGCAGTAAATTTGCGTATTTAAAAACCTATTCGTGCAACTT	1209
Db	1616	CTTTATTAATAAATTTGGAAGCAGTAAATTTGCGTATTTAAAAACCTATTCGTGCAACTT	1675
Qy	1210	AAATTTATATGATGTAATAAATGATCTTTTAGAGCTAGTAGGTACCCATAATGCTCAATA	1269
Db	1676	AAATTTATATGATGTAATAAATGATCTTTTAGAGCTAGTAGGTACCCATAATGCTCAATA	1735
Qy	1270	GGCACAGATCCAAATAGGAGTATTTTAATTGCAAGCAACGTGTACTTTAATCAATTTTAAA	1329
Db	1736	GGCACAGATCCAAATAGGAGTATTTTAATTGCAAGCAACGTGTACTTTAATCAATTTTAAA	1795

1301 CTGGGCTGGTTGGTACCCACCAACGGTCAGATCGGTAACGACCGAACCTGACATCCCTC 1360

...TAAATTTAGGATGTGATTGGTAC 1356

1237 TTAGAGCTAGTACGATCCATTAATGCTCAATAGCAACGATCCAAATAGCATATATTA 1296
1301 CTGGGCTGCTGGTGTGACCCCAACGATGCTGTAACGACCCGACCGGATGACATCTG 1360
1297 ATTGCAAGCAACGATGCTGATCTTAATCATTTAAAGATTAATTTTAAAGTGTGATGCTAC 1356
1361 ATGCTTCTTACGCTGATCTTCAACGACCTGAAAGCAAAATCTGCTGCTGACCTGATC 1420
1357 TTGTACTACATGAGGATGAGGATGACAAATGA 1388
1421 TTGCTTCCGACGATGAGGTTGACCAACGA 1452

RESULT 10
AA097492
ID AA097492 standard; DNA; 3769 BP.
XX
AC AA097492;
XX
XX 08-DEC-1995 (first entry)
XX
DE Plasmid pTECH2.
XX
XX htra promoter; vaccine; attenuation; vector; tetanus toxin;
KM pHTRA2; pTECH2; ds.
XX
XX Not specified.
OS
PN WO9520665-A.
XX
PD 03-AUG-1995.
XX
PE 31-JAN-1995; 95WO-GB00196.
XX
PR 31-JAN-1994; 94GB-0001795.
PA (MEDE-) MEDIVA HOLDINGS BV.
XX
PI Chatfield SN, Khan MA, LI J;
XX
DR WPI: 1995-275450/36.
XX
PT New DNA constructs confg. the htra promoter - used partic. for inducing
XX heterologous protein prodn. in vivo, esp. in vaccine compositions
XX
XX Disclosure; Page 31-34; pp: English.
XX
CC Vector plasmid pHTRA1 contains the S. typhimurium temp.-inducible
CC htra promoter sequence linked to the gene coding for tetanus toxin
CC fragment C (Tet-C). It was used to prepare pHTRA2 in which a
CC hinge region is present at the C-terminal of the Tet-C fragment.
CC This hinge region is obld. from pTECH2.
XX
XX Sequence 3769 BP; 982 A; 980 C; 863 G; 944 T; 0 other;

Query Match 57.2%; Score 796.8; DB 16; Length 3769;
Best Local Similarity 74.3%; Pred. No. 4.9e-128;
Matches 1005; Conservative 0; Mismatches 347; Indels 0; Gaps 0;

37 AAAAATCTGATTTGGGTGATTAATGAAGAAGATATAGATATATTAATAAAGAGT 96
101 AAAAACCTGATTTGGGTGACACAGCAAGACATGATGTTATCTCGAATAAGTCT 160
97 ACAATTTAAATTTAGATATTAATATATATATATATATATATATATATATATTA 156
161 ACCATTTCTGAATCTGACATCAACACATATTTATCTCGACATCTCTGTTCAACTCC 220
157 TCTGTAATATACATATCCAGATGCTCAATTTGTCGCCGGAATTAATAGCAAGCAATACAT 216
221 TCTGTTATACATATCCAGATGCTCAATTTGTCGCCGGAATTAATAGCAAGCAATACAT 280
217 TTAGTAAACATGATCTCTGAGATTATAGTACATAAAGCTATGATGATTAATATATAT 276

281 CTGGTTAACACGATCTTCTGAGTTATCTGCACAGGCCATGACATCGAATACAC 340
277 GATATGTTAAATTAATTTTACCGTCTTTGGTGGAGGTTCCCTTAAGTATCTGATG 336
341 GACATGTTCAACACCTTCAACCGTTACTCTGCTGGCGCTTCGAAAGTTTCTGCTCC 400
337 CATTTAGAACATATGCGCAACAAATGAGTATTCATTAATTAAGTCTATGAAACAACTAGT 396
401 CACCTGGAACAGTACGCGCACTAACGAGTACTCATCATCTGCTATGGAAGAAACACTCC 460
397 CTATCAATAGATCTGTTGAGTGTATCATCTTAAGGTAAATTAATATATATGAGCTTA 456
461 CTGTCATCTGCTGCTGTTGCTGCTGTTCCCTGAAGGTTACACCTGATCTGAGCTCG 520
457 AAAGATTCGCGGGAAGAACTAGACAAATTAACCTTTTAAAGGATTTTACCTGATTAATTAAT 516
521 AAAGATTCGCGGGAAGTTGCTGATCATCTTCCGACCTGCGGGAAGTTCAAC 580
517 GCTTATTAAGCAATTAATAGGTTTATTAATATTAATTAATTAATTAATTAATTAAT 576
581 GCTTATTAAGCAATTAATAGGTTTATTAATATTAATTAATTAATTAATTAATTAAT 640
577 AATTGTAATTAATTAATGAGTACTTATGGAAGTGCAGAAATTTCTGTTTAAAGGCTAT 636
641 AACCTGTACATCAACGGGCTTCTGATGGCTCCGCTGAATCATCTGCTGCGCTATC 700
637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
701 CGTAGGAGCAACATCACTCTTAAGGTGACCTTGCACCAACCAACCACTGCTGCTA 760
697 TCTATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 756
761 TCCATCGCAAGTTCCTGATCTTCTGCAAGGACGACGAAAGCAAGTCAAGAACTG 820
757 TACCAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 816
821 TATACAGCTACCTGCTATACACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
881 GACACCGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 940
877 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
941 ATCACTGATCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1000
937 TATTAATGAAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 996
1001 TACTACGAGCTGCTGATCAACAGGCTGAATTAATTAATTAATTAATTAATTAATTAAT 1060
997 GAATATGATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1056
1061 GAATATGATCTTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1120
1057 AATGAGCAATTTAGGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
1121 AAGCAACATCTGTTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1180
1117 CTAGAGTATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1176
1181 CTGAGTATGTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1240
1177 TTGCGGATTTAAACCTTATCTGTTCAACTTAAATTAATTAATTAATTAATTAATTAAT 1236
1241 CTGCGTACCTGAAACCTTCTGTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1300
1237 TTAGGACTAGTATGATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1296
1301 CTGGGCTGCTGTTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1360
1297 ATTGCAAGCAACGATGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356


```

Db 1241 CTGGGTGACCTGAAACCTACTCTGTTCAGCTGAAACTGTGACGACGACAAAACGCTTCT 1300
QY 1237 TTAGACTAGTAGTACCATTAATGTCAAATPAGCAACGATCCAAATAGGGATATATTA 1296
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1301 CTGGGTGTTGGTGTACCAACGACGATCGGTATCGATACGACCGATGATCTCTG 1360
QY 1297 ATGGCAAGCAACGTGCTTATATCATTTTAAAGATTTAAATTTTAGATGTGATGGTAC 1356
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1361 ATGGCTTTACAGTGTCTTCAACACCTGAAAGCAAAATCTGGCTTGGCACTGTAC 1420
QY 1357 TTTGTACCTACAGATGAGGATGACAAATGA 1388
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1421 TTGCTCCGACCGATGAGGTGGACCAACGA 1452

RESULT 12
AAO97491
ID AAO97491 standard; DNA; 1766 BP.
XX
AC AAO97491;
XX
DT 08-DEC-1995 (first entry)
XX
DE Plasmid pTECH2 + hinge region.
XX
KW htra promoter; vaccine; attenuation; vector; tetanus toxin;
KW pTECH2; pHR42; ds.
XX
OS Not specified.
XX
FH Key Location/Qualifiers
FT misc_feature 923..934
FT /tag= a
FT /function= hinge region
FT
XX WO9520665-A.
XX PN
XX PD 03-AUG-1995.
XX
PF 31-JAN-1995; 95WO-GB00196.
XX
PR 31-JAN-1994; 94GB-0001795.
XX
PA (MEDE-) MEDEVA HOLDINGS BV.
XX
PI Chatfield SN, Khan MA, Li J;
XX
DR WPI; 1995-275450/36.
XX
CC New DNA constructs contg. the htra promoter - used partic. for inducing
CC heterologous protein prodn. in vivo, esp. in vaccine compositions
CC
PS Disclosure: Page 34-35; pp; English.
XX
CC A 1.7 kb fragment (given in AAO97491) from pTECH2 and a 2 kb fragment
CC from pHR41 (AAO97490) were ligated to form pHR42, which incorporates
CC a temp.-inducible htra promoter from S. typhimurium operably linked
CC to the tetanus toxin G (Tet-C) fragment gene via a flexible hinge.
CC The construct was used to express Tet-C in attenuated Salmonella
CC for vaccine prodn.
XX
SQ Sequence 1766 BP; 434 A; 508 C; 419 G; 405 T; 0 other;
XX

Query Match 37.4%; Score 521; DB 16; Length 1766;
Best Local Similarity 72.9%; Pred. No. 9,3e-81;
Matches 671; Conservative 0; Mismatches 250; Indels 0; Gaps 0;
QY 468 GGGAGAAAGTTAGCAATATCTTTAGGGATTACCGTAATTAATCTATTTAGC 527
    || || || || || || || || || || || || || || || || || || || || || ||
Db 1 GGGCGAAGTTGCTCAGATACCTTTCCGCGACCTGCCGCAAGTTCAACGCTTACCTGGC 60
    || || || || || || || || || || || || || || || || || || || || || ||
QY 528 AATTAATGGTGTATTAATCTATTAATGATGATTAATCTTCTGCTAATTTGTATAT 587
    || || || || || || || || || || || || || || || || || || || || || ||

```

```

Db 61 TAACAATGGGTTTTTCATCACTATCACTAACGATCGTGTCTGTCTGTAACCTGTACAT 120
QY 588 AATGGAGTACTATATGGGAGTGCAGAAATTTACTGTTTAGACTATTAAGAGATTA 647
    || || || || || || || || || || || || || || || || || || || || || ||
Db 121 CAAGGCGGTTCTGATGGGCTCCCGTGAATCAGTGTGGGCTGCTACCTGTAGAGCA 180
QY 648 TAATTAACATTTAAACATAGATGATGATTAATTAATTAATTAATTAATTAATTAATTA 707
    || || || || || || || || || || || || || || || || || || || || || ||
Db 181 CAACATCACTCTTAAGCTGGACCGTTGCACACACACACACACACACACACACACACAC 240
QY 708 ATTAGATATTTTGCAGAACATTAATCCAAAAGACATTTGAAAATTTATACAGATTA 767
    || || || || || || || || || || || || || || || || || || || || || ||
Db 241 GTTCCGATCTTCTGCAAAACACTGAACCCGAAAGAGATGCAAAATCTGTATACAGCTA 300
QY 768 TTATCTATACCTTTTAAAGACCTCTGGGAAACCTTTAGCATATGATACAGATA 827
    || || || || || || || || || || || || || || || || || || || || || ||
Db 301 CCTGTCTATCACTCTCTCTGCTGATCTTTGGGTAAACCCGCTGGCTTACGACACAGATA 360
QY 828 TTATTTAATACAGTAGCTCTGTCTGTCTAAAGATGTTCAAATTTGAAAATTAACAGATTA 887
    || || || || || || || || || || || || || || || || || || || || || ||
Db 361 TTACCTGATCCGCGTAGCTCTAGCTCTAAAGAGCTTCAGCTGAAAACATCACGTGACTA 420
QY 888 TATGTTTGGACAAATGCGGCATCGTATACTAACGGAAATTTGATATATTTATAGAG 947
    || || || || || || || || || || || || || || || || || || || || || ||
Db 421 CATGTACCTGACCAACGCGCGCTCTACACTAACGGTAAACCTGTAACATCTACTACCGACG 480
QY 948 GTTATTAATGACATAAATTTATTAATAAAGATATACACTATATATGAAATAGATTTC 1007
    || || || || || || || || || || || || || || || || || || || || || ||
Db 481 TCTGTACAAAGCGCTGAATTTTCATCATCAACGCTACACCTCCGACACAGAAATGCTATTC 540
QY 1008 TTTTGTAAATCAGGTGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 1067
    || || || || || || || || || || || || || || || || || || || || || ||
Db 541 TTTGTTAAATCTGTGACTTCATCAAACTGTACGTTCTTACAAACAAACAAACAAACAT 600
QY 1068 TGTAGGTTATCCGAAAGATGGAATGCCCTTAATTAATCTTGATAGAAATTTAGAGTAG 1127
    || || || || || || || || || || || || || || || || || || || || || ||
Db 601 CGTGGTTACCCGAAAGAGGTACGCTTTCAACACCTGGACGAATTTGCGTGTGG 660
QY 1128 TTATATGCCCCAGATGATCCCTCTTTATATAAAGAGACAGTAAATTTGGTGATTT 1187
    || || || || || || || || || || || || || || || || || || || || || ||
Db 661 TTACAACGCTCCGGGTATCCCGGTATCAAAAATTTGAAGCTGTTAACCTGCTGACT 720
QY 1188 AAAAACTATTTGTGACAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1247
    || || || || || || || || || || || || || || || || || || || || || ||
Db 721 GAAAACTACTCTGTTCACCTGAACCTGTACGACGACGACGACGACGACGACGACGAC 780
QY 1248 AGGTACCCATATGTCATATGCAATGGCAACGATCCAAATPAGGATATATTAATTCACAA 1307
    || || || || || || || || || || || || || || || || || || || || || ||
Db 781 TGGTACCCACAAAGGTCAGATCGTACGACCCGACACCGTACATCTGTGCTGTAA 840
QY 1308 CTGTACTTTTAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1367
    || || || || || || || || || || || || || || || || || || || || || ||
Db 841 CTGTACTTTCAACACCTGAAAGACAAAATCTGGGTGGACATGTGATCTGTTCCGAC 900
QY 1368 AGATGAGGATGGAATAATGA 1388
    || || || || || || || || || || || || || || || || || || || || || ||
Db 901 CGATGAGGTTGGACCAACGA 921

RESULT 13
AAV30596
ID AAV30596 standard; DNA; 1535 BP.
XX
AC AAV30596;
XX
DT 07-DEC-1998 (first entry)
XX
DE Clostridium botulinum type G toxin C fragment gene in pETHsb.
XX
KW Antitoxin; vaccine; neurotoxin; toxin G; intoxication; immunogen;
KW botulism; BotG; ds.
XX
OS Clostridium botulinum serotype G 113/30 strain.

```

OS Synthetic.
 XX Key Location/Qualifiers
 FH CDS 108..1529
 FT /*tag= a
 XX
 XX W09808540-A1.
 XX
 XX PD 05-MAR-1998.
 XX
 XX PF 28-AUG-1997; 97WO-US15394.
 XX
 XX PR 28-AUG-1996; 96US-0704159.
 XX
 XX PA (OPHI-) OPHIDIAN PHARM INC.
 XX
 XX PI Thalley BS, Williams JA.
 XX
 XX DR WPI: 1998-230234/20.
 XX P-PSDB: AAW68400.
 XX
 PT Host cell containing recombinant expression vector encoding
 PT Clostridium botulinum type B or E toxin - useful to treat humans
 PT and other animals at risk of intoxication with clostridial toxin
 XX
 XX PS Example 49; Page 376-378; 428pp; English.
 XX
 CC This is the DNA sequence of the Clostridium botulinum serotype G
 CC (113/30 strain) neurotoxin fragment C gene contained in plasmid
 CC pETH15b. The encoded botC fragment C polypeptide (see AAW68400) has a
 CC His-tagged N-terminal extension. The vector can be used to express
 CC native (i.e., non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC C. botulinum toxins, especially type B and type E toxins. Methods
 CC are provided which allow for the isolation of soluble recombinant
 CC proteins free of significant endotoxin contamination. Preferred
 CC hosts for production of recombinant proteins are E. coli, insect
 CC cells and yeast cells. The recombinant toxins are used as
 CC immunogens for the production of vaccines and antitoxins that are
 CC useful in the treatment of humans and animals at risk of
 CC intoxication with clostridial toxin.
 CC
 XX Sequence 1535 BP; 596 A; 172 C; 246 G; 521 T; 0 other;
 XX
 SQ
 Query Match 17.8%; Score 248.2; DB 19; Length 1535;
 Best Local Similarity 55.9%; Pred. No. 5.5e-34;
 Matches 547; Conservative 0; Mismatches 408; Indels 24; Gaps 3;

DB 566 AATAATATGACATTAAATGATGTTAATGCAAAATCTTAATATATTTTTCGAATATAG 625
 OY 501 ACCGTAAATATTTAATGCTATTTAGCAAAATGAATGGCTTTTAACTATTAATAATGA 560
 DB 626 TATTAAGATATATATATGATATATATATATATATATATATATATATATATATATAT 665
 OY 561 TAGATATATCT 620
 DB 686 TAGATATGCTAAGCAAT 745
 OY 621 TCGTTAGAGCTATTAAGAGGATATATATATATATATATATATATATATATATATAT 680
 DB 746 AAACCTTAGATGAATTAATATCTAGATATATATATATATATATATATATATATATAT 805
 OY 681 TAAATATCAATACCTTTCTATGATATATATATATATATATATATATATATATATATAT 740
 DB 806 TACTACTAATATTTGTTGGATTAAGGATTTTATATATATATATATATATATATATATAT 865
 OY 741 AGAGATTGAAATATATACACAGTTATATATATATATATATATATATATATATATATAT 800
 DB 866 AGAAGTATCTCTCACTAT 925
 OY 801 AAACCTTACGAT 860
 DB 926 GAATCCTTTAATAGATGATACACAACTATCTGTTTAAATCAAG-----GTATGCAAAA 979
 OY 861 TGTTCATTTGAAATATATATACAGTTATATATATATATATATATATATATATATATAT 920
 DB 980 TATCTAT 1039
 OY 921 CGGAAATATGAT 980
 DB 1040 TAAATATGACGCAAT 1099
 OY 981 ATATACCTAT 1040
 DB 1100 AGCATCAATATCTCGAAT 1159
 OY 1041 TGTATCAT 1100
 DB 1160 TCTTAT 1178
 RESULT 14
 AAV30581
 ID AAV30581 standard; DNA; 1547 BP.
 XX
 AC AAV30581;
 XX
 DT 07-DEC-1998 (first entry)
 XX
 DE Clostridium botulinum toxin B fragment C gene in pETH15b.
 XX
 KW Antitoxin; vaccine; neurotoxin; toxin B; intoxication; immunogen;
 KW botulism; BoB; ds.
 XX
 OS Clostridium botulinum serotype B Danish strain.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS 108..1526
 FT /*tag= a
 XX
 XX W09808540-A1.
 XX
 XX PD 05-MAR-1998.
 XX
 XX PF 28-AUG-1997; 97WO-US15394.
 XX
 XX PR 28-AUG-1996; 96US-0704159.
 XX
 XX PA (OPHI-) OPHIDIAN PHARM INC.

XX Thalleys BS, Williams JA;
 PI WPI: 1998-230234/20.
 DR P-PSDB; AAM68394.
 XX Host cell containing recombinant expression vector encoding
 PT Clostridium botulinum type B or E toxin - useful to treat humans
 and other animals at risk of intoxication with clostridial toxin
 XX
 PS Example 35; Page 303-305; 428pp; English.
 CC This is the DNA sequence of the Clostridium botulinum serotype B
 CC (Danish strain) toxin fragment C gene contained in plasmid pETH15b.
 CC The encoded fragment C polypeptide (see AAM68394) has a His-tagged
 CC N-terminal extension. The vector was used to express native
 CC (i.e. non-fusion) soluble C fragment in Escherichia coli host
 CC cells. The invention relates to recombinant proteins derived from
 CC C. botulinum toxins. Methods are provided which allow for the
 CC isolation of soluble recombinant proteins free of significant
 CC endotoxin contamination. Preferred hosts for production of
 CC recombinant proteins are E. coli, insect cells and yeast cells.
 CC The recombinant toxins are used as immunogens for the production
 CC of vaccines and antitoxins that are useful in the treatment of
 CC humans and animals at risk of intoxication with clostridial toxin.
 CC
 SO Sequence 1547 BP; 634 A; 148 C; 263 G; 502 T; 0 other;
 Query Match 16.9%; Score 234.8; DB 19; Length 1547;
 Best Local Similarity 52.4%; Pred. No. 1.1e-31;
 Matches 689; Conservative 0; Mismatches 577; Indels 48; Gaps 6;

DB 806 ATTTATTTGGTGAATATTTTCACATATTTTAAATACGAATTTAACTCAATCAAAATATGA 865
 OY 750 AAAATTTACACAGATTAATTTATCTATTAACCTTTTAAAGACCTTCGGGAAACCTTT 809
 DB 866 AGAAGATATTAATTAATCAATCATATACGAAATTTTAAAGATTTTGGGAAATCTTT 925
 OY 810 ACATATGATACAGAAATTTATTTAAATACAGTACGCTTCTAGTCTAAAGATTTCAAT 869
 DB 926 AATGTACAAATTAAGAAATTTATATGTTTAAATCGCGGAATTAATATCATATTTAACT 985
 OY 870 GAAATATATACAGATTAATATGTTTGAACAATTCGCCATCTATCTACTACGAAAT 929
 DB 986 AAGAAAGATTCACCTCTAGCTGAAATTTTACACGATACGAAATTAATCAAAATCTTA 1045
 OY 930 GAATATATATTAATGAAGTTATATATATATGACATAATTTTAAAGATATACA-- 987
 DB 1046 ATATATTAATTAATGAAGTTATATATATTTGAGAAATTTATTTAAGAAAGAACTCAA 1105
 OY 988 -CCTATATATGAATATGATCTTTGTTAAATCGATGATTTTAAATATATATATATC 1046
 DB 1106 TTCTCATCTATTAATGATATATATATATGAGAAAGATTTATATATATATATATAT 1165
 OY 1047 ATATACAAATATATGACACATTTGATGTTATCCGAAGAT-----GGAATGCTTTAA 1100
 DB 1166 TATTTAAATCAAGAGTGGAGATATATATATATATATATATATATATATATATATAT 1225
 OY 1101 TATCTGATAGATATCTAAGATAGCTTAAATGCCCCAGTATCCCTTTTAAATA 1160
 DB 1226 ATTTGTTTACCTCTATATATATATATCTGATGATTTTACAAATCTATACAAATTAAGA 1285
 OY 1161 AATGAGCAGATAAATTTGCGATTTTAAACCTATCTGATCACTTAAATATATATGA 1220
 DB 1286 ATATGATGAGAACCCACATATATATGTTGCTGCTTTTAAAGATATGAAGATATC 1345
 OY 1221 TGATTAATAATGATCTTTAGAGCTAGTAGTACCAATATATGTCATATAGCAAGATCC 1280
 DB 1346 TGATGAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1405
 OY 1281 AATAGGATATATATATATGCAAGACAGTCTTAAATCAATTTAA----- 1329
 DB 1406 GTATTAAGATTTATTTTGTATTAAGTAATGTAATGTAATGTAATGTAATGTAATGTA 1465
 OY 1330 -GATTAATTTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1382
 DB 1466 TATTTAAATTTGGGATGTAATTTGGCAGTTTATCTTAAGATGAGAGGTGGAC 1519

RESULT 15
 AAV30579
 ID AAV30579 standard; DNA; 3876 BP.
 AC AAV30579;
 AC
 DT 07-DEC-1998 (first entry)
 XX
 DE Clostridium botulinum type B toxin gene from Danish strain.
 XX
 KW Antitoxin; vaccine; neurotoxin; toxin B; intoxication; immunogen;
 KM botulism; BotB; ds.
 OS Clostridium botulinum serotype B Danish strain.
 XX
 PN W09808540-A1.
 XX
 PD 05-MAR-1998.
 XX
 PF 28-AUG-1997; 97W0-US15394.
 XX
 PR 28-AUG-1996; 96US-0704159.
 XX
 PA (OPHI-) OPHIDIAN PHARM INC.

PI Thalley BS, Williams JA:

XX WPI: 1998-230234/20.

DR P-PSDB: AAM68392.

XX Host cell containing recombinant expression vector encoding
PT Clostridium botulinum type B or E toxin - useful to treat humans
PT and other animals at risk of intoxication with clostridial toxin

XX Example 35: Page 291-296; 428pp: English.

CC This is the coding region of the Clostridium botulinum serotype B
CC (Danish strain) toxin gene that codes for a 1291-amino acid
CC polypeptide (see AAM68392). The C fragment (see AAM68394) of the
CC toxin has been expressed as histidine-tagged protein in Escherichia
CC coli host cells. The invention relates to C. botulinum recombinant
CC toxin polypeptides. Methods are provided which allow for the
CC isolation of soluble recombinant proteins free of significant
CC endotoxin contamination. Preferred hosts for production of the
CC recombinant proteins are E. coli, insect cells and yeast cells.
CC The recombinant proteins are used as immunogens for the production
CC of vaccines and antitoxins that are useful in the treatment of
CC humans and animals at risk of intoxication with clostridial toxin.

XX Sequence 3876 BP; 1612 A; 370 C; 617 G; 1277 T; 0 other:

Query Match 16.88; Score 233.2; DB 19; Length 3876;

Best Local Similarity 52.4%; Pred. No. 2,1e-31;

Matches 688; Conservative 0; Mismatches 578; Indels 48; Gaps 6;

QY 90 AAAGATGACATTTTAAATTTAGATTTAATATGATTTATATACATATATCGGTT 149
DB 2583 AAATATATTTATCTTAATTTAGATTTAAGATTTAATATTTATATATATACATAT 2642
QY 150 TAATTCATCTGTAAATTAATCCAGATGCTCAATGCTGCCGGAATTAATGCCAAGC 209
DB 2643 TCGCGCAAGATGAGATGATGATGAGATGAGATGAGATGAGATGAGATGAGATG 2702
QY 210 AATTCATTTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 269
DB 2703 AACTGATGACCAAA-----TAGTAAATGATGAGATGATGATGATGATGATGAT 2753
QY 270 AATATATGATTTGTTTAAATTTTACGCTTTGCTTTGTTGAGGTTCTCTAAAGATC 329
DB 2754 ATTTAATAGTGTGCTTCTGTTTAAATTTAGGTTAGCTTTGATTAAGATTAATAT 2813
QY 330 TGTAGTCAATTTAACAATATGACAAATGATGATTTCAATATTTAGCTATGAAAA 389
DB 2814 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2873
QY 390 ACATAGTCTATCATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 449
DB 2874 TAAT-----TCGGCTGGAATATATCTTTAGGGTATATAGATTAATATG 2918
QY 450 GACTTTAAAGATTCGCGGAGAGATTAGACAAATTAACCTTTAGGGATTTAAGCTGAT 509
DB 2919 GACTTTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 2978
QY 510 ATTTAATGCTTATTTAGATTAATGATTAATGATTAATGATTAATGATTAATGAT 569
DB 2979 AGATTTATCAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 3035
QY 570 TTCTGCTAATTTGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 629
DB 3036 TAAGCTTAAATTTATTTATTTAGTAACTGATTAATGATTAATGATTAATGATTAATG 3095
QY 630 AGCTATTGAGAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 689
DB 3096 AGAGATTTATGCTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3155
QY 690 ATAGCTTTCTATTTAGATTAATTTAGATTAATTTAGATTAATTTAGATTAATTTAG 749
DB 3156 ATTTATTTGATTAATTTAGATTAATTTAGATTAATTTAGATTAATTTAGATTAATTTAG 3215

QY 750 AAATTTATACAGATTTATTTATCTATTAACCTTTTAAAGACTTCGGGAAACCTTT 809
DB 3216 AGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3275
QY 810 AGCATATGATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 869
DB 3276 AATGTACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3335
QY 870 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 929
DB 3336 AAAGAAAGATTTACCGTGTAGTGAATTTTAAACGATTAATTAATTAATTAATTAAT 3395
QY 930 GATTTATATTTAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 987
DB 3396 ATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3455
QY 988 -CCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1046
DB 3456 TTCTCAATCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3515
QY 1047 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1100
DB 3516 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3575
QY 1101 TAATCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1160
DB 3576 ATTTGTTTGTGCTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3635
QY 1161 AATGAGACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1220
DB 3636 ATATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3695
QY 1221 TGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1280
DB 3696 TGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3755
QY 1281 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1329
DB 3756 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3815
QY 1330 -GATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1382
DB 3816 TAATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3869

Search completed: December 1, 2002, 09:03:43
Job time : 270 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:58:27 ; Search time 56 seconds
(Without alignments)
7623.108 Million cell updates/sec

Title: US-09-816-467-3

Perfect score: 1392
Sequence: 1 atgggttttcaacacccaat.....aggatgacaaatgatata 1392

Scoring table: IDENTITY_MNC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1379.8	99.1	1858	1	US-08-668-381A-6
2	1346.4	96.7	1359	1	US-07-618-312A-1
3	1346.4	96.7	1359	1	US-08-280-228-1
4	1287.2	92.5	1359	1	US-08-110-786A-7
5	799.2	57.4	1359	1	US-07-618-312A-3
6	799.2	57.4	1359	1	US-08-280-228-3
7	228.2	16.4	3891	1	US-08-480-604A-27
8	228.2	16.4	3891	4	US-08-480-604A-27
9	228.2	16.4	3891	4	US-08-915-136-27
10	220.8	15.9	3509	4	US-09-255-829-19
11	83.4	6.0	161	1	US-08-110-786A-6
12	80	5.7	161	1	US-08-110-786A-4
13	75.6	5.4	1511	1	US-07-991-867B-8
14	75.6	5.4	1511	1	US-08-107-755A-8
15	75.6	5.4	1511	2	US-08-544-332-8
16	75.6	5.4	1511	4	US-09-370-861A-8
17	74.2	5.3	1330	1	US-08-480-604A-22
18	74.2	5.3	1330	2	US-08-405-496A-22
19	74.2	5.3	1330	4	US-08-915-136-22
20	74.2	5.3	1402	1	US-08-480-604A-25
21	74.2	5.3	1402	2	US-08-405-496A-25
22	74.2	5.3	1402	4	US-07-991-867B-32
23	73.8	5.3	660	1	US-08-107-755A-32
24	73.8	5.3	660	1	US-08-544-332-32
25	73.8	5.3	660	4	US-09-370-861A-32
26	73.8	5.3	660	4	US-08-852-629-11
27	73.8	5.3	4810	3	US-08-852-629-11

28	73.8	5.3	4838	3	US-08-852-629-15	Sequence 15, Appl
29	65.4	4.7	19124	2	US-08-487-826B-13	Sequence 13, Appl
30	62.4	4.5	3701	4	US-08-845-258-10	Sequence 10, Appl
31	62.4	4.5	3701	4	US-08-990-571-10	Sequence 10, Appl
32	62.4	4.5	3701	4	US-08-723-142A-10	Sequence 10, Appl
33	62.4	4.5	3701	4	US-09-528-784A-10	Sequence 10, Appl
34	61.8	4.4	19124	2	US-08-487-826B-13	Sequence 10, Appl
35	60.6	4.4	854	4	US-08-998-416-534	Sequence 534, App
36	60.6	4.4	854	4	US-08-998-416-186	Sequence 287, App
37	60.4	4.3	615	4	US-08-998-416-186	Sequence 186, App
38	59.8	4.3	663	4	US-08-998-416-191	Sequence 191, App
39	59.2	4.3	168575	4	US-08-929-329-1	Sequence 1, Appl
40	58.8	4.2	5433	3	US-08-307-499-36	Sequence 36, Appl
41	58.6	4.2	1323	1	US-08-307-499-1	Sequence 1, Appl
42	58.6	4.2	1323	1	US-09-299-268-36	Sequence 36, Appl
43	58.6	4.2	14176	1	US-08-307-499-1	Sequence 14, Appl
44	58.6	4.2	14176	1	US-08-307-499-14	Sequence 1, Appl
45	58.6	4.2	14176	4	US-09-299-268-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-668-381A-6
; Sequence 6, Application US/08668381A
; Patent No. 5780024
; GENERAL INFORMATION:
; APPLICANT: Brown, Robert H.
; APPLICANT: Fishman, Paul S.
; APPLICANT: Francis, Jonathan W.
; APPLICANT: Hosler, Betsy A.
; TITLE OF INVENTION: SUPEROXIDE DISMUTASE/TETANUS TOXIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/668,381A
; FILING DATE: 21-JUN-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,473
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,164
; REFERENCE/DOCKET NUMBER: 00786/263001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1858 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-668-381A-6
Query Match 99.1%; Score 1379.8; DB 1; Length 1858;
Best Local Similarity 99.9%; Pred. No. 7.6e-248;

MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Clostridium tetani
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1356
 US-07-618-312A-1

Query Match 96.7%; Score 1346.4; DB 1; Length 1359;
 Best Local Similarity 99.6%; Pred. No. 1.2e-241;
 Matches 1350; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 37 AAAATCTGGATTTGGTGTGATATGAGAAGATATGATATATTAATAAAGAGT 96
   |||
Db 4 AAAATCTGGATTTGGTGTGATATGAGAAGATATGATATATTAATAAAGAGT 63
QY 97 ACAATTTAAATAGATATTAATATATATATATATATATATATATATATATAT 156
   |||
Db 64 ACAATTTAAATAGATATTAATATATATATATATATATATATATATATATAT 123
QY 157 TCTGTATTAACATATCCAGATGCTCAATTTGGTCCCGAATAAATGGCAACATACAT 216
   |||
Db 124 TCTGTATTAACATATCCAGATGCTCAATTTGGTCCCGAATAAATGGCAACATACAT 183
QY 217 TTAGTAACATATGATCTTGCAGATATATGATATATATATATATATATATAT 276
   |||
Db 184 TTAGTAACATATGATCTTGCAGATATATGATATATATATATATATATATAT 243
QY 277 GATATGTTTAAATATTTTACCGTTAGCTTTTGGTGGAGGGTCCCAAGTATTCCTCTAGT 336
   |||
Db 244 GATATGTTTAAATATTTTACCGTTAGCTTTTGGTGGAGGGTCCCAAGTATTCCTCTAGT 303
QY 337 CATTTAGAACATATATGCGACAAATGATATTAATTAATTAATTAATTAATTAAT 396
   |||
Db 304 CATTTAGAACATATATGCGACAAATGATATTAATTAATTAATTAATTAATTAAT 363
QY 397 CTATCAATAGATCTGCTGGTGGAGGTGTCTACTTAAGGTATTAATTAATTAATTAAT 456
   |||
Db 364 CTATCAATAGATCTGCTGGTGGAGGTGTCTACTTAAGGTATTAATTAATTAATTAAT 423
QY 457 AAGATTCGCGGGGAGAGATTTAGACAATAATTAATTAATTAATTAATTAATTAAT 516
   |||
Db 424 AAGATTCGCGGGGAGAGATTTAGACAATAATTAATTAATTAATTAATTAATTAAT 483
QY 517 GCTTATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 576
   |||
Db 484 GCTTATTTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 543
QY 577 AATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 636
   |||
Db 544 AATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 603
QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
   |||
Db 604 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
QY 697 TCTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
   |||
Db 664 TCTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
QY 757 TAGACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
   |||
Db 724 TAGACAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 783
QY 817 GATACAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
   |||
Db 784 GATACAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
QY 877 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
   |||
Db 844 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 903

```

```

QY 937 TATATAGAGCTTATATATATGAGCTAAATTTATATATATATATATATATATATATAT 996
   |||
Db 904 TATATAGAGCTTATATATATGAGCTAAATTTATATATATATATATATATATATATATAT 963
QY 997 GAAATAGATCTTTGTTTAAATTCAGTGATTTTATTAATTAATTAATTAATTAATTAATTAAT 1056
   |||
Db 964 GAAATAGATCTTTGTTTAAATTCAGTGATTTTATTAATTAATTAATTAATTAATTAATTAAT 1023
QY 1057 AATGAGACATTTGAGTATATCCGAAGATGGAATGCCCTTAATTAATTAATTAATTAATTAAT 1116
   |||
Db 1024 AATGAGACATTTGAGTATATCCGAAGATGGAATGCCCTTAATTAATTAATTAATTAATTAAT 1083
QY 1117 CTAAAGATAGCTTATATATGCCCCAGGTATCCCTCTTATTAATTAATTAATTAATTAATTAAT 1176
   |||
Db 1084 CTAAAGATAGCTTATATATGCCCCAGGTATCCCTCTTATTAATTAATTAATTAATTAATTAAT 1143
QY 1177 TTCCGCTGATTTAAACCCATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAATTAAT 1236
   |||
Db 1144 TTCCGCTGATTTAAACCCATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAATTAAT 1203
QY 1237 TTAGGACTGTAGTACCCCAATATAGTCAATATAGGCAACGATCCCAATATAGGATATATTA 1296
   |||
Db 1204 TTAGGACTGTAGTACCCCAATATAGTCAATATAGGCAACGATCCCAATATAGGATATATTA 1263
QY 1297 ATTGCAAGCAACTGTACTTTAATCATTTTAAAGATTAATTAATTAATTAATTAATTAATTAAT 1356
   |||
Db 1264 ATTGCAAGCAACTGTACTTTAATCATTTTAAAGATTAATTAATTAATTAATTAATTAATTAAT 1323
QY 1357 TTTGTACTTAACAGATGAGGATGACAAATGATTA 1392
   |||
Db 1324 TTTGTACTTAACAGATGAGGATGACAAATGATTA 1359

```

RESULT 3
 US-08-280-228-1
 ; Sequence 1, Application US/08280228
 ; Patent No. 5571694
 ; GENERAL INFORMATION:
 ; APPLICANT: Makoff Dr, Andrew J
 ; APPLICANT: Romanos Dr, Michael A
 ; APPLICANT: Clare Dr, Jeffrey J
 ; APPLICANT: Fairweather Dr, Neil F
 ; TITLE OF INVENTION: VACCINES
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHVE P.C.
 ; STREET: 1100 No. 5571694th Gledbe Road
 ; CITY: Arlington,
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22201-4714
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/280.228
 ; FILING DATE: 25-JUL-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/618.312
 ; FILING DATE: 27-NOV-1990
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 8926832.0
 ; FILING DATE: 28-NOV-1989
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9006097.1
 ; FILING DATE: 17-MAR-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wilson, Mary J.
 ; REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 117-163
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 TELEX: 200797 NIXN UR
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1359 base pairs
 type: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Clostridium tetani
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1356
 US-08-280-228-1

Query Match 96.7%; Score 1346.4; DB 1; Length 1359;
 Best Local Similarity 99.6%; Pred. No. 1.2e-241;
 Matches 1350; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 37 AAAAATCGATGTTGGTGGTGAATGAAGAAGATATGATGTTATATTAATAAAGAGT 96
 DB 4 AAAAATCGATGTTGGTGGTGAATGAAGAAGATATGATGTTATATTAATAAAGAGT 63
 QY 97 ACAATTTAAATTTAGATATTAATGAATATATATATATATATATATATATATATATAT 156
 DB 64 ACAATTTAAATTTAGATATTAATGAATATATATATATATATATATATATATATATAT 123
 QY 157 TCTGTATTAATATCCAGATGCTCATGTCGCGCGGAATTAATGCAAAAGCAATTCAT 216
 DB 124 TCTGTATTAATATCCAGATGCTCATGTCGCGCGGAATTAATGCAAAAGCAATTCAT 183
 QY 217 TTACTAACAATGATGATGTTGAATATAGTGCATTAAGCTATGATATTTGAATATAT 276
 DB 184 TTACTAACAATGATGATGTTGAATATAGTGCATTAAGCTATGATATTTGAATATAT 243
 QY 277 GATATGTTTAATATTTTACCGTTAGCTTTTGGTTGAGGGTCTCTAAGATATCTGCTAGT 336
 DB 244 GATATGTTTAATATTTTACCGTTAGCTTTTGGTTGAGGGTCTCTAAGATATCTGCTAGT 303
 QY 337 CATTAGACATATGCGACAATGAGTATTCATTAATTTAGCTTATGAATAAATAGT 396
 DB 304 CATTAGACATATGCGACAATGAGTATTCATTAATTTAGCTTATGAATAAATAGT 363
 QY 397 CTATCAATAGATCTGTTGGAGTGTATCCTTAAGGTAATTAATTAATATATATATAT 456
 DB 364 CTATCAATAGATCTGTTGGAGTGTATCCTTAAGGTAATTAATTAATATATATATAT 423
 QY 457 AAAGATTCGCGGGAAGATTAAGCAAAATACCTTTTGGGATTTACCGTAAATTAAT 516
 DB 424 AAAGATTCGCGGGAAGATTAAGCAAAATACCTTTTGGGATTTACCGTAAATTAAT 483
 QY 517 GCTTATTAGCAAAATATGSGTTTATATACCTTATTAATGATATGATATGATATGAT 576
 DB 484 GCTTATTAGCAAAATATGSGTTTATATACCTTATTAATGATATGATATGATATGAT 543
 QY 577 AATTTGATATTAATGAGTACTTAATGGAAGTGCAGAAATTTACTGGTTTGGAGCTATT 636
 DB 544 AATTTGATATTAATGAGTACTTAATGGAAGTGCAGAAATTTACTGGTTTGGAGCTATT 603
 QY 637 AGAGGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
 DB 604 AGAGGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 663
 QY 697 TCTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 756
 DB 664 TCTATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 723
 QY 757 TACGACATGTTATTTATCTAATACCTTTTAAAGAGCTTCTGCGGGAACCCCTTTCAGATAT 816

DB 724 TACGACATGTTATTTATCTAATACCTTTTAAAGAGCTTCTGCGGGAACCCCTTACGATAT 783
 QY 817 GATPACAGATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
 DB 784 GATPACAGATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
 QY 877 ATPACAGATATTTATTTATTTGACAAATGCGCCATCGTATCTAAGCGAAATTAATTAAT 936
 DB 844 ATPACAGATATTTATTTATTTGACAAATGCGCCATCGTATCTAAGCGAAATTAATTAAT 903
 QY 937 TATTTAAGAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
 DB 904 TATTTAAGAGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 963
 QY 997 GAATATATCTTTTGTTAATCAGTGATTTATTAATTAATTAATTAATTAATTAATTAAT 1056
 DB 964 GAATATATCTTTTGTTAATCAGTGATTTATTAATTAATTAATTAATTAATTAATTAAT 1023
 QY 1057 ATGAGCAGATTTAGTGTATTCGCAAGATGCAAAATGCCCTTATTAATTAATTAATTAAT 1116
 DB 1024 ATGAGCAGATTTAGTGTATTCGCAAGATGCAAAATGCCCTTATTAATTAATTAATTAAT 1083
 QY 1117 CTAAAGTAGTTATTAATGCCCCAGTATCCCTTTATTAATTAATTAATTAATTAATTAAT 1176
 DB 1084 CTAAAGTAGTTATTAATGCCCCAGTATCCCTTTATTAATTAATTAATTAATTAATTAAT 1143
 QY 1177 TTGGGATTTTAAAAACCTATCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1236
 DB 1144 TTGGGATTTTAAAAACCTATCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1203
 QY 1237 TTAGACTAGTAGTACCATTAATGTCAAATAGCAACGATGCAAAATAGGAGATATTA 1296
 DB 1204 TTAGACTAGTAGTACCATTAATGTCAAATAGCAACGATGCAAAATAGGAGATATTA 1263
 QY 1297 ATTGCAAGCACTGTACTTTATCATTTTAAAGTAAATTTAAGATGATGATGATGAT 1356
 DB 1264 ATTGCAAGCACTGTACTTTATCATTTTAAAGTAAATTTAAGATGATGATGATGAT 1323
 QY 1357 TTGTACCTACAGATGAGGATGAGCAAAATGATTA 1392
 DB 1324 TTGTACCTACAGATGAGGATGAGCAAAATGATTA 1359

RESULT 4
 US-08-110-786A-7
 : Sequence 7, Application US/08110786A
 : Patent No. 5443966
 : GENERAL INFORMATION:
 : APPLICANT: FAIRWEATHER, Neil Fraser
 : APPLICANT: MAKOFF, Andrew Joseph
 : TITLE OF INVENTION: Expression of tetanus toxin fragment C
 : NUMBER OF SEQUENCES: 8
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Nixon & Vanderhye P.C.
 : STREET: 1100 No. 5443966th Glebe Road
 : CITY: Arlington
 : STATE: Virginia
 : COUNTRY: U.S.A.
 : ZIP: 22201-4714
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/110,786A
 : FILING DATE: 23-AUG-1993 1991
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 07/777,337
 : FILING DATE: 29-NOV-1991
 : PRIOR APPLICATION DATA:


```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/618,312A
FILING DATE: 19910516
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
APPLICATION NUMBER: 510-51
TELECOMMUNICATION INFORMATION:
TELEPHONE: 0101 703 8750400
TELEFAX: 0101 703 5253468
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1359 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Clostridium tetani
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1356
US-07-618-312A-3

Query March 57.4%; Score 799.2; DB 1; Length 1359;
Best Local Similarity 74.3%; Pred. No. 2.9e-140;
Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

QY 37 AAAAACTGGATTGGTGGTGATATGCAAGAAATATGTTATTTAAAAAGAGT 96
DB 4 AAAAACTGGATTGGTGGTGATATGCAAGAAATATGTTATTTAAAAAGAGT 63
QY 97 ACAATTTAAATTTGATATTAATATGATTTATATCAGATATATCTGGGTTAAATCA 156
DB 64 ACCATTCTGAACTTGGACATACACAGCATATATCTCCGACATCTCGGTTCAACTCC 123
QY 157 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGGAATTAATGCAAGCAATACAT 216
DB 124 TCTGTAATACATATCCAGATGCTCAATTTGGTCCCGGAATTAATGCAAGCAATACAT 183
QY 217 TTAGTAACATGATCTTCTGGAATATAGTCAATAAGCAATGATGATATATAT 276
DB 184 CTGGTTAAACGAATCTTCTGGAATATAGTCAATAAGCAATGATGATATATAT 243
QY 277 GATTGTTAAATTTAAATTTAACTGTTGTTGAGGGTTCTTAAAGTATCTGTA 336
DB 244 GACATGTTCAACAACTTCACTGCTTCTGCTGCGGTTCCGAAGTTCTGCTGCC 303
QY 337 CATTTAGAACAAATATGCAACAATAGTATTCATTAATTAAGCTCTATGAAAAACATAGT 396
DB 304 CACCTGGAACAGTACGCGCTAAGAGTACTCATCATCATGCTATGAAAGAACACTCC 363
QY 397 CATATCAATGATCTGCTGTTGATGATCACTTAAAGCAATTAATTAATTAATTAATTA 456
DB 364 CTGTCCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 423
QY 457 AAAGATTTCCGCGAGAGAGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 516
DB 424 AAAGATTTCCGCGAGAGAGTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 483
QY 517 GCTTATTTAGCAAAATTAATGAGGTTTATTAATTAATTAATTAATTAATTAATTAATTAAT 576

```

```

DB 484 GCCTACCTGGCTAACAAATGGTTTTTCATCATCATCATCAACGATCGTCTTCTGCT 543
QY 577 AATTTGATATTAATAGAGTACTTATGGGAAGTGCAGAAATTAATGCTGTTAGAGCTATT 636
DB 544 AACCTGATCATCAACGCGCTTCTGATGGCTCCGCTGAAATCACTGCTGCGGCGCTATC 603
QY 637 AGAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 696
DB 604 CGTGAGGACAAACAATCATCTTTAAAGCTGACCTGACCTGACCTGACCTGACCTGACCTG 663
QY 697 TCTATTTGATTAATTTAGGATTTTTCGAAAGCAATTAATTAATTAATTAATTAATTAATTA 756
DB 664 TCCATGCAAGAGTCCCTATCTCTGCAAGCACTGAAACCGAAGAGATGGAAGAACTG 723
QY 757 TACACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 816
DB 724 TATACAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 783
QY 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 876
DB 784 GACACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 843
QY 877 ATAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 936
DB 844 ATCACTGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 903
QY 937 TATATTAAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 996
DB 904 TACTACGAGCTCTGTTACAAAGCGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 963
QY 997 GAAATAGATTTCTTTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1056
DB 964 GAAATAGATTTCTTTGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1023
QY 1057 AATGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1116
DB 1024 AACGAAACATGTTGTTACCAAGAGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1083
QY 1117 CTAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1176
DB 1084 CTGCGTGTGTTACCAAGCGCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1143
QY 1177 TTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1236
DB 1144 CTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1203
QY 1237 TTAGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1296
DB 1204 CTGCGTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1263
QY 1297 ATTGCAAGCACTGTTACTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1356
DB 1264 ATGCTTTCAATGTTACTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1323
QY 1357 TTTGTTACCTACGATGAGGATGCAAAATGATTA 1392
DB 1324 TTGCTTCCGACGATGAAGTTGACCAACGACTAA 1359

RESULT 6
US-08-280-228-3
Sequence 3, Application US/08280228
Patent No. 5571694
GENERAL INFORMATION:
APPLICANT: Makoff Dr, Andrew J
APPLICANT: Romanos Dr, Michael A
APPLICANT: Clare Dr, Jeffrey J
APPLICANT: Fairweather Dr, Neil F
TITLE OF INVENTION: VACCINES
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P. C.
STREET: 1100 No. 5571694th Glebe Road

```

CITY: Arlington,
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patenlin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/280,228
FILING DATE: 25-JUL-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/618,312
FILING DATE: 27-NOV-1990
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 8926832.0
FILING DATE: 28-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9006097.1
FILING DATE: 17-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mary J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 117-163
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
TELEX: 200797 NIXN UR
INFORMATION FOR SEQ. ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1359 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Clostridium tetani
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1356
US-08-280-228-3

Query Match 57.48; Score 799.2; DB 1; Length 1359;
Best Local Similarity 74.3%; Pred. No. 2,9e-140;
Matches 1008; Conservative 0; Mismatches 348; Indels 0; Gaps 0;

37 AAAAACTCGATTGGTGGGTGATATGAAGAAGATATAGTCTTATTTAAAGAGT 96
4 AAAAACCTGATTGTTGGTGGACACGAAGAAGACATCGATGTTATCCGAAAAAGTCT 63
97 ACAATTTAAATTTAGATATTAATATGATATTAATTCAGATATATCGGTTAAATTC 156
64 ACCATTCTGAACCTGGACATCAACAACGATATATATCCGACATCTCTGTTTCAACTCC 123
157 TCTGTAATTAACATATCCGATGCTCAATTTGGTCCCGGAATAAATGCAAGCAATACAT 216
124 TCTGTAATTAACATATCCGATGCTCAATTTGGTCCCGGATCAACGCAAGCTATCCAC 183
217 TTAGTAACAATGAGTCTGAGTATAGTATAGCAATTAAGTATGATATTTGAATTAAT 276
184 CTGTTAAACAGATCTTCTTAAGTATATCGTCAACAGGCAATGACATCGAATTCACAC 243
277 GATATGTTAATTAATTTACGTTAGCTTTGGTTCAGGGTCTCAAAATATCTGCTAGT 336
244 GACATGTTCAACAACTTCACGTTAGCTTTGGTTCAGGGTCTCCGAAATTTCTGCTTCC 303
337 CATTTAGACAAATATGGCAAAATGAGTATTCATTAATTAAGCTATGAAAAACATACT 396

Db 304 CACCTGGACAGTACGCGACCTAACGAGTACATCCATCAGCTCTATGAGAAACACTCC 363
QY 397 CTATCAATAGGATCGTGGTTGGATGATATCACTTAAGGAATTAATTAATGAGCTTAA 456
Db 364 CTGTCATCGGCTCTGGTGGTCTGTTTCCCTGAAAGGATACAACTGATCTGGACTGCT 423
QY 457 AAGATTTCCCGGGGAGAGATTAGCAAAATTAATTTAGGATTTACCTGATTAATTAAT 516
Db 424 AAGACTCCCGGGGAGAGATTAGCAAAATTAATTTAGGATTTACCTGATTAATTAAT 483
QY 517 GCTTATTTAGCAATTAATGAGTGGTTTATTAATTAATTAATTAATTAATTAATTAAT 576
Db 484 GGTACCTGCTTAACAAATGAGTGGTTTATTAATTAATTAATTAATTAATTAATTAAT 543
QY 577 AATTGATTAATTAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 636
Db 544 AACCTGATCAATCAACGCGTTCGATGAGGCTCCGCTGAATCAGCTGCTGCTGCTGCT 603
QY 637 AGAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 696
Db 604 CGTAGAGACACAACTCACTTTAAGCTGACCGTTCGCAACCAACCAACCACTACCTA 663
QY 697 TCTATTGATAAATTTAGATATTTTGAAGCATTAATCCAAAGAGATTGAAGAAATTA 756
Db 664 TCCATGACAGATTCGCTATCTCTGCAAGCACTGAACCCGAAAGATGGAAGAACTG 723
QY 757 TACACAGTATTTATCTATTAACCTTTTAAGACACTTGGGGAAACCCCTTACGATAT 816
Db 724 TATACAGCTACCTCTATCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
QY 817 GATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 876
Db 784 GACACGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 843
QY 877 ATACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 936
Db 844 ATCACTGATCACTATGATCACTGACCAACGCGCTCCACACTAAGCTTAACCTAATC 903
QY 937 TATTAATGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 996
Db 904 TACTACGACGCTGTGACACGCGCTGAATTAATTAATTAATTAATTAATTAATTAAT 963
QY 997 GAATAGATTTCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1056
Db 964 GAATCGATTTCTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023
QY 1057 AATGACATTTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1116
Db 1024 AAGCAACATCTGTTGGTATCCCGGAAGCGGTAACTGTTCAACCAACCTGACAGATT 1083
QY 1117 CTAAGAGTAGGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1176
Db 1084 CTGGGTGTTGGTTACACGCTCCGCTGATCCCGCTGATCAAAAAATTAATTAATTAATTAAT 1143
QY 1177 TTGCTGATTTAAAACTATTTCTGTACAACTTAATTAATTAATTAATTAATTAATTAAT 1236
Db 1144 CTGGGTGATGAAACCTATCTGTGTACAGTGAAGTGTAGACACAAAAAGCTTCT 1203
QY 1237 TTAGGACTAGTAGGATACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1296
Db 1204 CTGGGTGATGAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1263
QY 1297 ATTGACAGCACTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1356
Db 1264 ATGCTCTTAACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1323
QY 1357 TTGCTACCTACAGATGAGGATGAGCAAAATGATTA 1392
Db 1324 TTGCTTCGACCGATGAAGTTGAGACCAACGACTAA 1359

RESULT 7
US-08-480-604A-27


```

APPLICATION NUMBER: US/08/915,136
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/480,604
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/405,496
FILING DATE: 16-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/329,154
FILING DATE: 25-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/161,907
FILING DATE: 02-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/985,321
FILING DATE: 04-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/429,791
FILING DATE: 31-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: INGOLIA, DIANE E.
REGISTRATION NUMBER: 40,027
REFERENCE/DOCKET NUMBER: OPD-01763
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 705-8410
TELEFAX: (415) 397-8338
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3891 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3888
US-08-915-136-27

```

```

Query Match 16.4% Score 228.2; DB 4; Length 3891;
Best Local Similarity 55.9%; Pred. No. 2e-34;
Matches 556; Conservative 0; Mismatches 403; Indels 36; Gaps 5;

QY 86 TAAAGAGTACATTTTAAATTTAGATATTAATGATATTAATCAGATATATCTG 145
DB 2618 TTATTAATACCTCATATGATTTAAGATATGAAGATATTAATGATATATCTA 2677
QY 146 GGTTAATTCATCTGTAATTAACATATCCAGATGCTCAATGGTCCCGGAATTAATGCA 205
DB 2678 GGTATGATCAAAATAATATATGTTAGTAAAGTAAATTTTGATCC--AATAGATAAAA 2734
QY 206 AAGCAATGATTTAGTAACAATGATCTTCGAGTTATAGTATGATTAAGATATGATA 265
DB 2735 ATCAAAATTCATTTATTTATTTAGAACTAGTAAATGAGTAAATTTTAAAAATGCTA 2794
QY 266 TTGAATATGATATGATTTTAAATTTTACCGTTAGCTTTTGGTGAAGCTTCCAAAG 325
DB 2795 TTGATATATATGATGATGATGAATTTTGTAGTACGTTTGGATGAAGATTCCTAAGT 2854
QY 326 TATCTGCTAGCATTTAGAACATATGCGACAAATGAGATTCATTAATTTAGCTCATGA 385
DB 2855 ATTTTAAACGT-----ATAACTCTAAATTAATGAATATATCAATTAATTAATTTGATGG 2905
QY 386 AAAAACAATGATCTATCATAGATGATGCTGATGATGATCACTTAAAGTATTAATCTTAA 445
DB 2906 AAAATTAAT-----TCAGATGGAAGTATCACTTAAATTTATGTTGTAATTA 2950
QY 446 TATGAGACTTTAAAGATTCGCGGAGAGATTTAGACAAATAACTTTTATAGGATTTACCTG 505
DB 2951 TCTGAGCTTTACAGGATCTCTCGAATAAATAAAGAGATGATTTTAAATACAGTCANA 3010
QY 506 ATAATTTAATGCTTATTTAGCAAAATAATGGGTTTATTAATCTATTTACTAATGATAGAT 565

```

```

DB 3011 TGATTAATATATCGATTTATATTAACAGATGATTTTGTAACTATCATCTAATATAGAT 3070
QY 566 TATCTCTGCTAATTTGATATTAATGAGTACTTATGGGAAGTGAATTTCTGCTG 625
DB 3071 TAAATACCTCTAAATTTATTAATTAAGAGATTAATGATCAAAAACCAATTTCAATTT 3130
QY 626 TAGACCTATTAGAGAGATTAATTAATTAACATTAACATTAACATTAATTAATTAATTA 685
DB 3131 TAGTAATATTTCTAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3190
QY 686 ATCAATACGTTCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 745
DB 3191 ATGATATATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3250
QY 746 TTGAAAAATTTATACAGATTTATTTATCTTAACCTTTTATAGAGCTTTGGGAAC 805
DB 3251 TCAAGATTTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3310
QY 806 CTTTACGATATGATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 865
DB 3311 ATTTACATATGATTAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3370
QY 866 AATTGAAAA-----TATACAGATTTATGATTTGACAAATGCGCATCTATACTA 919
DB 3371 ATGTAATTAATGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3430
QY 920 ACGGAATTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 979
DB 3431 TGACTACAAACATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3490
QY 980 GATATACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1039
DB 3491 AATATGCTCTGGAAT--AAAGATATATATTTAGTAATTAATTAATTAATTAATTAATTA 3547
QY 1040 ATGATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1074
DB 3548 ATGATAGTTTAAATAAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3582

RESULT 10
US-09-255-829-19
; Sequence 19, Application US/09255829
; Patent No. 6461617
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN, & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,829
; FILING DATE: 23-FEB-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB97/02273
; FILING DATE: 22-AUG-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/782,893
; FILING DATE: 27-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ESMOND, ROBERT W.

```


Db	3516	TAATTAAATCAAGAGTGGAGAGATATACCTTTAAATATTTTAAAGAAAGACGAAGAAAA	35757
Oy	1101	TAATCTCATGAGATTTCTAAGAGTACGTTATTAATGCCCCAGGATGCCCTTTATATAAAA	1160
Db	3576	ATTCTGTTTTAGCTCTTAATAGATTTCTATAGTGTTTTGAATACCTATACAAATATAAAGA	36355
Oy	1161	AATGGAACGATMAAATTTGCGATTTTAAAAACATTCCTGTACAACTTAAATTTATAGA	1220
Db	3636	ATATGATGAACAGCCAAACATATATSTTGCAGTTGCTTTTAAAAAAGATGAAAGAAATGAC	3695
Oy	1221	TGATAAAAAATGCACCTTTAGGACTACTAGCTACCAATATGTCAAATATGCAACGATCC	1280
Db	3696	TGATGAGATATGAGTTGATTTGATATAGTAAAGTAAAGGTAATTTGAAGA	3755
Oy	1281	AAATAGGATATATTTAATTCACGACAACTGTACTTTATCATTTAAAA-----	1329
Db	3756	GTAATAAGATTATTTTGTATAAAGTAAAGTAGTAAAAAGGCTAAAAAAGAAACATA	3815
Oy	1330	GATAAATTTTGGATGCGATGTGTACTTTACTTCCTCAGATGGGGATGAC	1382
Db	3816	TAATTTTAAATTTGGAGTAAATTTGGCAGTTTATTCCTTAAATATGAAGGCTGAC	3869

```

1 RESULT 3
2 US-10-051-952-2
3 : Sequence 2, Application US/10051952
4 : Patent No. US20020107199A1
5 : GENERAL INFORMATION:
6 : APPLICANT: Walker, Patricia
7 : TITLE OF INVENTION: Methods of Administering Botulinum Toxin
8 : FILE REFERENCE: 2933CIP
9 : CURRENT APPLICATION NUMBER: US/10/051,952
10 : CURRENT FILING DATE: 2002-01-17
11 : PRIOR APPLICATION NUMBER: 09/730,237
12 : PRIOR FILING DATE: 2000-12-05
13 : NUMBER OF SEQ ID NOS: 12
14 : SOFTWARE: PatentIn Ver. 2.1
15 : SEQ ID NO 2
16 : LENGTH: 3876
17 : TYPE: DNA
18 : ORGANISM: Botulinum toxin
19 :
20 US-10-051-952-2

```

Query Match	16.68;	Score 231.2;	DB 12;	Length 3876;
Best Local Similarity	52.7%;	Pred. No. 1.3e-37;		
Matches 693;	Conservative 0;	Mismatches 573;	Indels 48;	Gaps 7

Db	2583	AAATTAATTAATCTTAAATTAAGATATAGATATAAATTAATTAAATGATTATCAGATATCTGGGTT	149
QY	150	TAATTCATCTGTATTAACATATCCAGATGCTCATTTGGTGGCCGGAATTAATGGCAAGC	209
Db	2643	TGGAGCAAGSTAGAGCTATA-----TGTATGGGGTCAAGCTTAATGATTAATAATCA	2693
QY	210	AATACATTTAGTAACAAATCAATCTTCTGAGCTTAATAGCATTAACGTAATGATATGA	269
Db	2694	ATTTAAATTAAGTACTAGTCCACACAGATAGTAAGATTAGAGTCACTCAAAATCAGAAATTAAT	2753
QY	270	ATATAAGATATGTTTAATTAATTTAATCCGTTAGCTTTGGTGAAGGTTCCATAAGTATAC	329
Db	2754	ATTTAAATAGATATGTCCTGATTTTGGCTGTAGCTTTTGATTAAGATATCCATAATATAG	2813
QY	330	TGCTAGCTATTGGAACAATATGGCCAAATAGATATTCATTAATTAATAGCTGTATGAAAA	389
Db	2814	GAAATGATGATATCAAAAATTAATTTTCATATATATAAATATCGAATTAATTAATGTATGAAAA	2873
QY	390	ACATTAAGTATCAATAGATCTGGTGGAGCTATACACTTAAGGTAAAGGTAACTAATATATG	449
Db	2874	TAAT-----TCAGGCTGGAAATATCTTTATAGGGTAAATAGATTAATATG	2918
QY	450	GACTTTAAAGATTCGGCGGAGAAAGTTAGACAATAATCTTTAGGGATTTAAGCTGATTA	509

Db	2919	GACCTTAATGTTAATAAATGGAAAAACCAATCAGATATTTTTCATCTAAACATAGAGA	2978
Oy	510	ATTATACCTATTTAGCAAAATAATGGCGTTTTTATAACTATTAAGTAGATAGATTTC	569
Db	2979	AGATATACACAGCTATATAAATAGATGGTTTTTGTAACTATTCTAATAAT - - -TTGGA	3035
Oy	570	TTTCGCTAATTTGTATATAAATGGAGTACTATGGAAGTCACAGAAATTAAGCTTTAGG	629
Db	3036	TAAATGCTAAATTTATATTAATTAATGCGACGCTTGGAATCAAAATATGCAATTAAGATATAG	3095
Oy	630	AGCTATTAAGAGAGATTAATTAATTAACATTAACCTAGATGATTAATTAATTAATCA	689
Db	3096	AGAAGTATATGTTAATGCTGAATAATACATTTAAATTGATGCTATAGTAGTAACACA	3155
Oy	690	ATAGCTTCTATATATAATTAGAGATTTTTCGAAAGCATTAATCCMAAAGATTTGA	749
Db	3156	ATTATATTGGATGAATATTTTAACTATTTTAAATACCAATTAATCAATCAATATTAA	3215
Oy	750	AAAAATTAACAGAGTTATTTATCTATTAACCTTTTAAAGAGACTTGGGGAAACCCCTT	809
Db	3216	AGAGATATTAATAAATCAATCATATATACGAAATAGTTAAAGATTTTGGGGAATCCTTT	3275

```

Y 3276 AATGATATAAATAAGAAATTTATATATGTTTAATGCGGGGAATAAAAATTCATATATTAACT 3335
D 3276 AATGATATAAATAAGAAATTTATATATGTTTAATGCGGGGAATAAAAATTCATATATTAACT 3335
Y 870 GAAATAATATAACGATATATATGATTTGGACAATTCGCCCATCGTATATCTAAGCAAAATT 929
D 3336 AGGAAAGATTCCTCTCTAGGTGAAAATTTAAATACGTAGCAAAATTAATTCAGATTTCCAA 3395
Y 930 GAATATATATATATAGAAGTTATATATATGACTATAAATTTATATATAAAGATATACA-- 987
D 3396 TATATATAAATATAGAAATTTATATATATGAGAAAATTTATATATATAAGAGAGCTCAAA 3455
Y 988 -CCTAATAATGAATAGATTCCTTTGTTTAATCAGGTATTTTAAATATATATGTATC 1046
D 3456 TTCTCATCTATATAAGATGATATAGTTAGAAAAGAAAGATTTATATCATCTAGATTTGGT 3515
Y 1047 AATATACAATAATGAGCAGATTTAGTTATCCGAAAGATGGAAATG---CCTTAATAA 1103
D 3516 ACTTCCACCTGAAAGAGTGGAGATATATCCCTATAAATTTTAAAGAACAGAGCAAAA 3575
Y 1104 TCTGTATATAATTTCTAAGAGTATGTTATATATGCCCCAGGTATCCCTTTATATAAAAAT 1163
D 3576 ATTTGTTTTATCTATATATATAGAGATTCCTATAAATTTTAAAGACTATATGAATAAAGA 3635
Y 1164 GGAAGCAGTAA--ATTCCGATTTTAAACCACTATCTGTACAACTTAATTTATAGA 1220
D 3636 AATGATGTAAGCAGCATCATATAGTTGTACAGTTGCTTTTAAAAAAGATGAAGAAAGTAC 3695
Y 1221 TGATATAAATGCATCTTTAGAGCTAGTAGAATCCCATATGCTCAAAATAGCAACGATCC 1280
D 3696 TGATATATAGATTTGATTTGATTCATCTGTTTCTAGAAATCGGAGTTTACGTAAAAA 3755
Y 1281 AAATGGGATATATTAATTCGACGAACATGCGACTTTATCATTTAAAAAGTAA----- 1334
D 3756 GATATAAATATATTTTGTATATAGTAAATGTGATTTAAAAAGAGTAAAAAGAAACCATATA 3815
Y 1335 -----AATTTTAGATGTGATTTGTTACTTTGTACTCTACAGATGAGGATGAC 1382
D 3816 TAAGTCAAAATTTGGATGTAAATTTGGCACTTATTCCTAAACATGAGAGGTGGAC 3869

RESULT 4
US-10-051-952-1
; Sequence 1, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; CURRENT FILING DATE: 2002-01-17

```

; PRIOR APPLICATION NUMBER: 09/730,237
 ; PRIOR FILING DATE: 2000-12-05
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: Patentl Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 3891
 ; TYPE: DNA
 ; ORGANISM: botulinum toxin
 ; US-10-051-952-1

Query Match 16.4% Score 228.2; DB 12; Length 3891;
 Best Local Similarity 55.9%; Pred. No. 4.4e-32;
 Matches 556; Conservative 0; Mismatches 403; Indels 36; Gaps 5;

Oy 86 TAAAGAGTACATTTTAAATTTAGATATTTAAATGATATTAATCAGATATCTG 145
 Db 2618 TTATTAACTCTTAAATGATATTTAGATATTTAGAAAGTAACTATTAATACCTATCTA 2677
 Oy 146 GGTAAATTCATCTGTATTAACATATCCAGATGCTCAATGTGCCCCGAGTAAATGCGA 205
 Db 2678 GGTATGATCAAAAATAAATTTGGTAGTAAAGTAAATTTGATCC--AATAGATAA 2734
 Oy 206 AAGCATATCATTATAGTAAACATGATCTCTGAAGTTATAGTCAATAAAGCTATGATA 265
 Db 2735 ATCAATATCATTTATTTAATTTAGAAAGTAAATTTGAGTAAATTTTAAATAATGCTA 2794
 Oy 266 TTGAATTAATGATATGTTTAAATTTTACCCTTACCTTTGGTTGAGGGTCTTAAG 325
 Db 2795 TTGTATTAATAGTATGTATGATGAAATTTAGTACGCTTTGGATTAAGAAATCTTAAGT 2854
 Oy 326 TATCGTAGTATTTAGAACAAATGGCAATAGATGATCAATTAATTTAGCTCTATGA 385
 Db 2855 ATTTTAACAGT-----ATAGCTTAATATGATATATCATTAATTAATTTGATGG 2905
 Oy 386 AAAACATAGTCTATCAATAGAGTCTGTTGGAGTATCAGTTAAAGSTAAATTAATCTTA 445
 Db 2906 AAAATTAAT-----TCAGATGGAAGATTCACCTTAATTAATGCTGAATTA 2950
 Oy 446 TATGACCTTAAAGATTCGGGAGAGATGAGCAAAATTAATCTTTAGGATTTACTG 505
 Db 2951 TCTGACCTTACAGATCTCAGAAATTAACAAAGAGTACTTTTAAATAGAGTCAAA 3010
 Oy 506 ATAAATTTATGCTTATTTAGCAAAATTAATGGCTTTTAACTATTTACTAATGATAGT 565
 Db 3011 TGATTAATATATCAGATTTATTAATGAAACAGATGATTTTGTACTATTAATTAATGAT 3070
 Oy 566 TATCTCTCTAATTTGTATTAATAGAGTCTTATGGAGTGCAGAAATTTACTGCT 625
 Db 3071 TAAATTAATCTTAAATTTTATTAATGAAAGATTAATTAATGAAACCAATTTCAATTT 3130
 Oy 626 TGGAGCTTTTGGAGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 685
 Db 3131 TAGGTATATTTCTAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3190
 Oy 686 ATCAATAGTCTTCTATGATTAATTTAGATTTTGCAGAGCTTAATCCAAAGAGA 745
 Db 3191 ATGATATATTTGATTAATTAATTTAATCTTTTGAATGAGATTAATTAATGAAAGAAA 3250
 Oy 746 TTGAAAAATTAATCAGAGTATTTATCTAATACCTTTTACAGACTTGGGGAAC 805
 Db 3251 TCAAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3310
 Oy 806 CTATGATATGATAGATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 865
 Db 3311 ATTTACATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3370
 Oy 866 AATTTGAAAA-----TATACAGATTAATTAATTTGACAAATGGCCCATCTGATATCTA 919
 Db 3371 ATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3430
 Oy 920 ACGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 979
 Db 3431 TGACTACAAACATTTATTTAATTTCAAGTTTGTATAGGGGACAAATTTATTAATAAAA 3490

Oy 980 GATATACCCAAATTAATGAAATAGATTTCTTTGTAATCAGTGAATTTATTAATTAAT 1039
 Db 3491 AATATGCTTCTGCAAA--AAAGATATATTTGTAGAAATTAATGATGCTATATATTA 3547
 Oy 1040 ATGTATCATTAATCAATTAATGAGCAGATTTGAGT 1074
 Db 3548 ATGTATGATTTAAATAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 3582

RESULT 5
 US-09-288-326-10
 ; Sequence 10, Application US/09288326
 ; Patent No. US20010018049A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kei Roger Noki
 ; APPLICANT: George Sachs
 ; TITLE OF INVENTION: Method and Compositions for the
 ; FILE REFERENCE: 17282
 ; CURRENT APPLICATION NUMBER: US/09/288,326
 ; CURRENT FILING DATE: 1999-04-08
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 4835
 ; TYPE: DNA
 ; ORGANISM: Clostridium botulinum
 ; US-09-288-326-10

Query Match 16.4% Score 228.2; DB 10; Length 4835;
 Best Local Similarity 55.9%; Pred. No. 4.5e-32;
 Matches 556; Conservative 0; Mismatches 403; Indels 36; Gaps 5;

Oy 86 TAAAGAGTACATTTTAAATTTAGATATTTAAATGATATTTAATCAGATATCTG 145
 Db 2975 TTATTAACTCTTAAATGATATTTAGATATTTAGAAAGTAACTATTAATAGCTATCTA 3034
 Oy 146 GGTAAATTCATCTGTATTAACATATCCAGATGCTCAATGTGCCCCGAGTAAATGCGA 205
 Db 3035 GGTATGATCAAAAATAAATTTGGTAGTAAAGTAAATTTGATCC--AATAGATAA 3091
 Oy 206 AAGCATATCATTATAGTAAACATGATCTCTGAAGTTATAGTCAATAAAGCTATGATA 265
 Db 3092 ATCAATATCATTTATTTAATTTAGAAAGTAAATTTGAGTAAATTTTAAATAATGCTA 3151
 Oy 266 TTGAATTAATGATATGTTTAAATTTTACCCTTACCTTTGGTTGAGGGTCTTAAG 325
 Db 3152 TTGTATTAATAGTATGTATGATGAAATTTAGTACGCTTTGGATTAAGAAATCTTAAGT 3211
 Oy 326 TATCTCTCTAATTTGTATTAATAGAGTCTTATGGAGTGCAGAAATTTACTGCT 385
 Db 3212 ATTTTAACAGT-----ATAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3262
 Oy 386 AAAACATAGTCTATCAATAGAGTCTGTTGGAGTATCAGTTAAAGSTAAATTAATCTTA 445
 Db 3263 AAAATTAAT-----TCAGATGGAAGATTCACCTTAATTAATGAGTAAATA 3307
 Oy 446 TATGACCTTAAAGATTCGGGAGAGATTAATGACAAATTAATTTAGGATTTTACTG 505
 Db 3308 TCTGACCTTACAGATCACTCAGAGAAATTAACAAAGAGATGTTTAAATTAATTAATTA 3367
 Oy 506 ATAAATTTATGCTTATTTTGCAGAAATTAATGGGTTTATTAATTAATTAATTAATTAAT 565
 Db 3368 TGATTAATTAATCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3427
 Oy 566 TATCTCTCTAATTTGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 625
 Db 3428 TAAATTAATCTAAATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3487
 Oy 626 TATGACCTTATTAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 685
 Db 3488 TAGGTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3547

EARLIER APPLICATION NUMBER: US 60/092,416
; EARLIER FILING DATE: 1998-07-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 9
; LENGTH: 1327
; TYPE: DNA
; ORGANISM: Clostridium botulinum
; FEATURE:
US-09-350-756-9

Query Match 15.3%; Score 213; DB 10; Length 1327;
Best Local Similarity 55.6%; Pred. No. 1,8e-29;
Matches 539; Conservative 0; Mismatches 400; Indels 46; Gaps 6;

86 TAAAGAGTACAAATTTTAAATTTAGATATTAATGATATTAATTCAGATATCTG 145
11
44 TTATTAATCTCTATATTTAGATATTAATGATATTAATTCAGATATCTG 103
146 GGTTAATCTCTATATTAATGATATTAATGATATTAATTCAGATATCTG 205
104 GGTATGATCAAAATTAATTTGATGATTAATTTGATCC--AATGATATAA 160
206 AAGCAATACATTTAGTAAACATGATCTTGAAGTATAGTCATTAAGCTATGATA 265
161 ATCAATTCATTTTATTAATTTAGATATTAATGATATTAATTCAGATATCTG 220
266 TTGATATTAATGATATTTATTAATTTAGATATTTAGATATTTAGATATTTAG 325
221 TTGATATTAATGATATTTATTAATTTAGATATTTAGATATTTAGATATTTAG 280
326 TATCGATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 385
281 ATTTTACAGATAT-----AAGCTAATTAATGATATTAATTCAGATATTTAG 331
386 AAAACATAGTCTATCAATGATGATCTGTTGAGTATCACTTAAGATATACTTAA 445
332 AAAATTAAT-----TCAGATGAAAGTATCACTTAATTAATGATATAA 376
446 TATGATCTTAAAGATTCGCGGAGAGATTAAGCAATTAATCTTTGGGATTTACCTG 505
377 TCTGATCTTAAAGATTCGCGGAGAGATTAAGCAATTAATCTTTGGGATTTACCTG 436
506 ATAAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 565
437 TGATTAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 496
566 TATCTTCTGATATTTGATATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 625
497 TAAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 556
626 TAGGATCTTAAAGATTCGCGGAGAGATTAAGCAATTAATCTTTGGGATTTACCTG 685
557 TAGGATCTTAAAGATTCGCGGAGAGATTAAGCAATTAATCTTTGGGATTTACCTG 616
686 ATCAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 745
617 ATGATTAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 676
746 TTGAAAAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 805
677 TCMAAGATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 736
806 CTTTACGATATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 865
737 ATTTACGATATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 796
866 AATGAAAAA-----TATACAGATTAATGATATTTAGATATTTAGATATTTAG 909
797 ATGATTAATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 856
910 TCGTATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 969

857 TGACTACAAATGACATCAAAATTTATTAATTCAGATTTGATAGGGGACAAAATTT 916
970 ATTTAAAGATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 1029
917 ATTTAAAGATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 973
1030 ATTTAAAGATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 1074
974 GTATATTTATTAATGATCTTATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 1018

RESULT 8

US-10-051-952-7
; Sequence 7, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 09/730,237
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 3759
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-7

Query Match 15.0%; Score 208.4; DB 12; Length 3759;
Best Local Similarity 53.7%; Pred. No. 1.3e-28;
Matches 571; Conservative 0; Mismatches 451; Indels 42; Gaps 5;

16 CCAATTCATTTCTTATTTCTTAAATCTGATGTTGGTATTAAGGAAGATATA 75
2470 CATTTAAGCTTTCTTATTTCTTAAATCTGATGTTGGTATTAAGGAAGATATA 2529
76 GATGTTATTTAAAGATTAATGATATTTAGATATTTAGATATTTAGATATTTAG 135
2530 TTTAAGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2589
136 GATATATCTGTTTAAATCTGATATTTAGATATTTAGATATTTAGATATTTAG 195
2590 GATATCTGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2646
196 ATAAATGCAAAATCAATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 255
2647 ACTAATTAATCAATTTAGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2706
256 GCTATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 315
2707 AATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2766
316 GTTCTAAAGTATCTGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 375
2767 ATTCCTAA-----CTATGATTAATGATATTTAGATATTTAGATATTTAG 2820
376 AGCTATGAAAAAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 435
2821 AATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2868
436 AATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 495
2869 AATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2928
496 GATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 555
2929 TATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 2988
556 AATGATTAATGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAGATATTTAG 615

Db 2989 AATGATACATTAGGAGATCTTAACCTTTATATATTAATGCAAAATTTAATAGATCAAAATATCA 3048
 ;
 QY 616 ATTAAGTGTGTTAGAGCTATTAGAGAGATTAATTAATTAACATTTAAAGTAGATGATG 675
 ;
 Db 3049 ATTTAAATTTAGTAATATCATGTTAGTGCATATATTTTAAATAGTATTTGT 3108
 ;
 QY 676 AATAATATTAATCAATAGCTTTCTATTTGATAAATTTAGATATTTTGGCAACATTAAT 735
 ;
 Db 3109 AGTATATCAAA--GATATATGTATGTAGATATTTTAAATTTTGTATGAATATAGAT 3165
 ;
 QY 736 CCAAAAGAGATTTGAAAAATTTATACAGATTTATTCATACCTTTTAAAGAGCTTC 795
 ;
 Db 3166 GAACAGAAATTCACACTTTTATATGCAATGACCTAATACAAATTTTGAAGGATTTT 3225
 ;
 QY 796 TGGGGAACCCCTTACGATATGATACAGAAATATTTAATACCAAGTACCTCTAGTCT 855
 ;
 Db 3226 TGGGGAATATTTTCTTTATGACAAAGATATCTTTATTTA----- 3267
 ;
 QY 856 AAAGATGTTCAATTTGAAAAATATATACAGATTAATATGATTTGACAAATGCGCATGTA 915
 ;
 Db 3268 AATGTGTTAAACCAAAATTAACCTTTATTTGATAGAGAAAAAGATTTCTAATTTAAGCATTTAT 3327
 ;
 QY 916 ACTAAGGAAAAATTTGAATATATTTATTTAGAGGTTATATTAATGACATAAATTTTATA 975
 ;
 Db 3328 AATTAAGAAAGCACTATTCTTTAGCTATATAGATTAATAGTGAATTAAGTTAAATA 3387
 ;
 QY 976 AAAAGATATACCTTAATTAATGAATAGATTTCTTTGTTAATCAGTGATTTTATTA 1035
 ;
 Db 3388 CAAGAATATTAATTAATAGTACTAGCAATATCTTTAGTAAAGAAATGATCAGTTATAT 3447
 ;
 QY 1036 TTATATGATCATTAATCAATTAATGAGACATTTGAGTTATCC 1079
 ;
 Db 3448 ATTAATTTTGTAGCCAGAAACCTACTATTTTCCATTATATGC 3491
 ;
 RESULT 9
 US-08-981-087A-5
 ; Sequence 5, Application US/08981087A
 ; Patent No. US20020081304A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Elmore, Michael J.
 ; APPLICANT: Mauchline, Margaret L.
 ; APPLICANT: Minton, Nigel P.
 ; APPLICANT: Pasechnik, Vladimir A.
 ; APPLICANT: Tilball, Richard W.
 ; TITLE OF INVENTION: TYPE F BOTULINUM TOXIN AND USE THEREOF
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NIXON & VANDERHYE P.C.
 ; STREET: 1100 NO. US20020081304A1th Giebe Rd. 8th floor
 ; CITY: Arlington
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22201-4741
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/981,087A
 ; FILING DATE: 27-MAY-1998
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/GB96/01409
 ; FILING DATE: 12-JUN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: GB 9511909.5
 ; FILING DATE: 12-JUN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Crawford, Arthur R.
 ; REGISTRATION NUMBER: 25,327
 ; REFERENCE/DOCKET NUMBER: 124-688

; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 703-816-4000
 ; TELEFAX: 703-816-4100
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1293 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; US-08-981-087A-5
 ;
 Query Match 14.8%; Score 206.2; DB 8; Length 1293;
 Best local Similarity 56.1%; Pred. No. 2.8e-28;
 Matches 551; Conservative 0; Mismatches 393; Indels 39; Gaps 7;
 ;
 QY 74 TAGATGTTATATTTAAAGAGTACAAATTTTAAATTTAGATATTAATGATATTAAT 133
 ;
 Db 44 TATATTAATAAATAAAGATACTCTATTTAGATATGCCATATGAATAATTAATTA 103
 ;
 QY 134 CAGATATATCTGGTTAATTCATCTGTAATACATATCCAGATGCTCAATTTGTCGCCG 193
 ;
 Db 104 TAGATATCTGTGATATGTTCAATATTAAGCATTAATGGAATGT--ATATATTTAT 160
 ;
 QY 194 GAATTAATGCAAGCAATACATTTAGTAAACAATGATCTTCTGAAGTTATAGCAT 253
 ;
 Db 161 CAAACAATAGAAATTCATTTGGAATATATAGTAAAGCCTAGTAACTTAATAGCTC 220
 ;
 QY 254 AAGCTATGATTTGAATTAATGAATATGTTAAATTTTAAATTTTAAATTTTAAATTT 313
 ;
 Db 221 AAATATATGATATTAATTAATCAATAGTATCAAAATTTTAAATTTTAAATTTTAA 280
 ;
 QY 314 GGGTCTTAAAGTATCTGCTAGTATTTAGAACATATGCGACAAATGAGATTTCAATA 373
 ;
 Db 281 GGAATTCCTAATTAATCTTCAATA-----AAGTGAATCTTAATTAATTAATTAAT 331
 ;
 QY 374 TTAGCTATAGAAAAAACAATAGTCTATCAATAGATCTGTTGAGTATCACTTAAG 433
 ;
 Db 332 TAGATGTTAATAGGAATTAAT-----TAGGATGGAATAATATCACTTAAT 379
 ;
 QY 434 GATATTAATTAATGAGCTTTAAAGATTTCCGGGAGAGATTAGCAATAACTTTTA 493
 ;
 Db 380 ATTAATAAATAAATTTGAGCTTTCAAGATCTGCGAAATAATCAAAACTAGTTT 439
 ;
 QY 494 GGAATTTACCTGATAATTAATTAATGCTATTTAGCAAAATAATGGTTTATATCTAT 553
 ;
 Db 440 ATTATACAAATGATTAATATATCTGATTAATTAATTAATGATTAATTAATTAAT 499
 ;
 QY 554 CTAAATGATGATTAATCTGCTAATTTTGTATTAATTAATGAGTACTATGGAAGTCA 613
 ;
 Db 500 CTAAATTAATGATTAAGCAATTTAGCAATTTTCAATCAATGAATTTAATGATGA 559
 ;
 QY 614 AATTAATGTTTAAAGGATTTAGAGAGATTAATTAATTAATTAATTAATTAATTAAT 673
 ;
 Db 560 CAATTTCAATTTAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 619
 ;
 QY 674 GTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 733
 ;
 Db 620 GTAAATGATCAAG--ATATGTTGATTAATTAATTAATTAATTAATTAATTAATTA 676
 ;
 QY 734 ATCCAAAGAGATTGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 793
 ;
 Db 677 GTAAACAGAAATTTAGACTTATATAGTATAGAGGACAGATCCAGATCTTAAAGACT 736
 ;
 QY 794 TCTGGGAACCCCTTACGATATGATACGAATATTAATTAATTAATTAATTAATTAAT 853
 ;
 Db 737 TTTGGGAATTAATTTGTTATTAATTAATTAATTAATTAATTAATTAATTAATTA 792
 ;
 QY 854 CTAAAGATGTTCAATTTGAAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 913
 ;
 Db 793 ACAGATTAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 852
 ;
 QY 914 ATACTAACGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 973

Best Local Similarity 51.7%; Pred. No. 7.8e-24;
Matches 520; Conservative 0; Mismatches 466; Indels 19; Gaps 4;

QY 16 CCAATTCATTTTCTTATTTAAATATCGGATTTGGGATTAATGCAAGAGATATA 75
Db 2524 CCATTTGAACCTTCTCTATATCTATATATTAATTTTATATATTTTATATGATTA 2583
QY 76 GATGTTATATTTAAAGAGTCAATTTTAAATTTAGATATTAATATGATTTATATCA 135
Db 2584 TATATAAAATTAAGATATTTTATTTAGATATGCGATATGAAATATATTAATTTATA 2643
QY 136 GATATATCTGGTTTAAATCTCTATATTAATATGATTTCCAGATCTCAATTTGGCCCGGA 195
Db 2644 GATATCTCTGATATGTTCAATATATAGCAATTAATGGAACGTATATTTATTC--A 2700
QY 196 ATAAATGCAAGCAATATCAATTTAGTAAACAATGCAATCTTGAAGTTATAGCATATA 255
Db 2701 ACAATATGAATCAATTTGATATATATATAGTAGCTTATGCAAGTTATATAGCTCA 2760
QY 256 GCTATGATTTGATTAATATATGATATGTTTATATTTTACCCTTATGCTTTGGTTGAG 315
Db 2761 ATATATGATATATATATGAAATAGATATCAAAATTTTATATATGTTTCTGGTAAAG 2820
QY 316 GTTCTTAAGATCTGATCTATTTAGCAATATGCAACAATGCAACAATGATTCATTAAT 375
Db 2821 ATTCCTAAA-----CACTAGAACCTATGATCATATTCGGGAATAGACTATA 2868
QY 376 AGCTCTTGAAGAAACATAGTCTATCAATAGGATCTGTTGGAGTGTATCACTTAAAGT 435
Db 2869 ATAAATTTGATGGGATTAATTAATTCGGGATGGAATATACATAGAGCTTTAGAGT 2928
QY 436 AATTAATTAATGATTTTAAAGATTTCCGGGAGAGATAGCAATATACCTTTAGG 495
Db 2929 TGTCAATTAATTTGGACTTTTCAAGATATCTTCTGAAATTAAGAAATTTTATTTTAGG 2988
QY 496 GATTTACCTGATTAATTAATGCTTATTTAGCAATATATGCTTTTATATCTATCT 555
Db 2989 TATCAAGACCTTATAGGATATCTAATTTATATATATATATGATTTTGTATATCT 3048
QY 556 AATGATGATTAATCTTCTGATTTTATATTAATTAATGAGTATTTAGGAAGTCAGAA 615
Db 3049 AATTAATGATTTAGGCAATTTAGAAATTTAGATCAATGAAATTTAATAGTGAATAATCA 3108
QY 616 AATTAATGATTTAGGATTTAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 675
Db 3109 ATTGCAATTTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAGGAT 3168
QY 676 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 735
Db 3169 GATGATGAA--CGTATGTTGATTAAGATTTTAAAGTTTAAATTTAGGATTTAGGAT 3225
QY 736 CCAAAAGAGTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 795
Db 3226 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3285
QY 796 TGGGGAACCTTTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAGGAT 855
Db 3286 TGGGGAATTTATTTGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3345
QY 856 AATGATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 914
Db 3346 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3405
QY 915 TACTATGCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 974
Db 3406 GGTCTGTTTATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3465
QY 975 AAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1019
Db 3466 GGTCTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3510

RESULT 12

US-10-051-952-5
; Sequence 5, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; PRIOR FILING DATE: 2002-01-17
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 3831
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-5

Query Match 11.2%; Score 155.8; DB 12; Length 3831;
Best Local Similarity 51.9%; Pred. No. 2.5e-19;
Matches 451; Conservative 0; Mismatches 382; Indels 36; Gaps 3;

QY 9 TTCAACACCAATTCATTTCTTATTTCTAAATATCTGATTTGGGTTGATTAATGAGA 68
Db 2514 TACAATGCCCTTTAATTTTTCATATATTAATTTCTTTATTAAGATATATTA 2573
QY 69 AGATATAGATCTTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 128
Db 2574 TGAATATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2633
QY 129 TATATCAGATATATCTGGGTTTATTTATTCATCTGTAATTAATCAATCAATGCTCAAT 188
Db 2634 TTATAGGATATCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2688
QY 189 GCCCGAATTAATGCAAAACATATCAATTTAGTAAACATGAAATCTTGAAGTTATAGT 248
Db 2689 -CTTAATGATATTAATCAATGATTTAATTAATTAATTAATTAATTAATTAATTAAT 2747
QY 249 GCATTAAGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 308
Db 2748 AATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2807
QY 309 GTTGAAGGTTCTTAAGATTTCTGCTAGTCAATTTAGAACATTAATGCAAAATGATATTC 368
Db 2808 GATTAAGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2852
QY 369 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 428
Db 2853 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2897
QY 429 TAAAGTAACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 488
Db 2898 TAGGATATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2957
QY 489 TTTTGAAGATTTTACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 548
Db 2958 TTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3017
QY 549 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 608
Db 3018 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3077
QY 609 TGCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 668
Db 3078 TCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3137
QY 669 TAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 728
Db 3138 TGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3197
QY 729 ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 788

Db 3198 ATTAAGTATGAAGATATTATATATATGAGCAAAATATTAAAGATCTTATTA 3257
QY 789 AGACTTCTGGGGAACCCCTTACGATATGATACGATATTTATTAATCCAGTACCTC 848
Db 3258 AGATTTATGGGAAATCCCTTGAAGTTTGATACGATATTTATTTAATGATATTA 3317
QY 849 TACTTCTAAAGATGTCATGTGAATA 877
Db 3318 TATGATAGCTATATTGACCTGAAAGTA 3346

RESULT 13
US-09-350-756-8
; Sequence 8, Application US/09350756
; Patent No. US2002004521A1
; GENERAL INFORMATION:
; APPLICANT: U.S. Army Medical Research Institute for Infectious Diseases
; APPLICANT: John S. Lee
; APPLICANT: Peter Pushko
; APPLICANT: Michael D. Parker
; APPLICANT: Jonathan F. Smith
; APPLICANT: Mark T. Dertzbaugh
; APPLICANT: Leonard Smith
; TITLE OF INVENTION: Botulinum Neurotoxin Vaccine
; FILE REFERENCE: 003/124/SAP RIID 98-21
; CURRENT APPLICATION NUMBER: US/09/350,756
; EARLIER FILING DATE: 1998-07-09
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Apple Macintosh Microsoft Word 6.0
; SEQ ID NO 8
; LENGTH: 1987
; TYPE: DNA
; ORGANISM: Clostridium botulinum
; FEATURE:
US-09-350-756-8

Query Match 10.4%; Score 145.2; DB 10; Length 1987;
Best Local Similarity 55.5%; Pred. No. 1.7e-17;
Matches 386; Conservative 0; Mismatches 273; Indels 37; Gaps 4;

QY 86 TAAAAAGAGTACATTTTAAATTTAGATATTAATTAATGATTTATTCAGATATATCTG 145
Db 1319 TTATTAATAGCTTATATTTGATTTAGATATGAAGTAATCATTTATAGACTTACTA 1378
QY 146 GGTATTAATGATCTGATTAATGATATCCAGATGCTCAATTTGTCGCCGGAATAATGGCA 205
Db 1379 GGTATGATCAAAAATAATATTTGGTAGTAAGTAATTTTGATCC--AATGATTAATA 1435
QY 206 AAGCAATACATTTAGTAACAAATGATCTCTGAAGTTATAGTCATTAAGCTATGGATA 265
Db 1436 ATCAATATTCATTTTATTAATTTAGAAAGTAAATTTAGGATATTTTAAAAATGCTA 1495
QY 266 TTGAATTAATGATATGTTAATTAATTTTACGTTACCTTTGGTTGAGGTCCTTAAG 325
Db 1496 TTGAATTAATGATATGTTAATTAATTTTACGTTACCTTTGGTTGAGGTCCTTAAGT 1555
QY 326 TATCTGCTGATTTAGAACAAATATGCGCAATGATGATTAATGATGCTCTATGA 385
Db 1556 ATTTTAAACAGT-----ATAGTCTAAATTAATGATATTAATTAATTTGATGG 1606
QY 386 AAAAACAATAGTCTATCAATAGATCTGTTGGAGTGTATCACTTAAGGTAATAACTTAA 445
Db 1607 AAAATTAAT-----TCAGATGGAAGTATCACTTAATTTATGCTGAATAATA 1651
QY 446 TATGACTTTTAAAGATTTCCGGGAGAGATTAGAACAAATTAATTTTATGAGATTTACTG 505
Db 1652 TCTGACTTTTACAGGATACAGGAAATTAACAAAGAGTGTTTTAAATATACGTCAAA 1711
QY 506 ATTAATTAATGCTTATTAATGATTAATGATTTTATTAATGATTAATGATTAATGAT 565
Db 1712 TGAATTAATATATGATTAATTAATTAACGATGATTTTGTAACTATCTAATATATAGAT 1771

QY 566 TATCTCTGCTAATTT-----TGATATTAATGAGTACTTATGGGAAGTCAGAA 615
Db 1772 TAAATAGCTTAATTAATCTTAATAATTTATATTAATGAAAGATTAATAGATCAAAAACA 1831
QY 616 ATTACTGTTTATGAGCTATTTAGAGAGATTAATTAATTAATTAATTAATTAATTAAT 675
Db 1832 ATTCAATATTTAGTAAATATTTATGCTAGTAATTAATTAATTAATTAATTAATTAAT 1891
QY 676 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 735
Db 1892 AGAGATACACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1951
QY 736 CCAAAAGATTTGAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 771
Db 1952 GAAAAAAGATTTCAAAAGATTTATTAATTAATTAATTAATTAATTAATTAATTAAT 1987

RESULT 14
US-10-051-952-4
; Sequence 4, Application US/10051952
; Patent No. US20020107199A1
; GENERAL INFORMATION:
; APPLICANT: Walker, Patricia
; TITLE OF INVENTION: Methods of Administering Botulinum Toxin
; FILE REFERENCE: 2933CIP
; CURRENT APPLICATION NUMBER: US/10/051,952
; PRIOR FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 09/730,237
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentlin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3876
; TYPE: DNA
; ORGANISM: botulinum toxin
US-10-051-952-4

Query Match 8.2%; Score 113.8; DB 12; Length 3876;
Best Local Similarity 49.7%; Pred. No. 6.5e-12;
Matches 481; Conservative 0; Mismatches 422; Indels 64; Gaps 5;

QY 9 TTCAACCAATTTCTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 68
Db 2526 TACAAATCCCTTAATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2585
QY 69 AGATATGATGTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 128
Db 2586 TGAATATTTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2645
QY 129 TATATCAGATATATCTGGTTTAAATGATCTGTAAATTAATTAATTAATTAATTAAT 188
Db 2646 TTTATGATGATCAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2705
QY 189 GCCCGGAT---AAATGCAAGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 245
Db 2706 TCCAAATATTTCAATTTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2765
QY 246 AGTCATTAAGCTATGATTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 305
Db 2766 AGTAAACCAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2825
QY 306 TTGTTGAGGTTCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 365
Db 2826 TTTGATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2857
QY 366 TTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 425
Db 2858 -----TACCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2912
QY 426 ACTTAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 485
Db 2913 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2972

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 08:57:42 ; Search time 1951 seconds

(without alignments)
11555.160 Million cell updates/sec

Title: US-09-816-467-3

Sequence: 1 atggttttttaacaccacaaat.....aggagtgacaaatgatataa 1392

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_estro:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_iny:*
20: em_gss_pln:*
21: em_gss_vtl:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_trod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101.2	7.3	1101	17	CNS00EVL
2	92.6	6.7	1101	17	AL069706 Drosophila
3	89.2	6.4	1190	17	CNS00396
4	89	6.4	1201	17	CNS020N7
5	88.4	6.4	1190	17	CNS0167M
6	87.4	6.3	1101	17	CNS00EVL
					AL069706 Drosophila

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
CNS00EVL												
7	87	6.2	1101	17	CNS003BD							
8	83.6	6.0	1101	17	CNS00EVL							
9	80.8	5.8	1001	17	CNS0155H							
10	80.2	5.8	876	17	CNS009GL							
11	79.4	5.7	829	17	CNS03LUB							
12	79.4	5.7	890	17	AQ026918							
13	78.6	5.6	1101	17	CNS00EPO							
14	78.2	5.6	1101	17	CNS0039G							
15	78.2	5.6	1101	17	CNS00EVL							
16	77.8	5.6	1169	17	CNS06KHO							
17	77.6	5.6	1101	17	CNS00EVL							
18	77.4	5.6	1027	17	CNS02750							
19	77.4	5.6	1027	17	CNS00EVL							
20	77.4	5.5	1101	17	CNS00088							
21	76.4	5.5	820	17	B11728							
22	76.2	5.5	897	10	BB620968							
23	75.6	5.5	1084	17	CNS071NH							
24	75	5.4	1101	17	CNS00KAE							
25	74.8	5.4	893	17	CNS013XE							
26	74.8	5.4	1092	12	BF274559							
27	74.6	5.4	1101	17	CNS020K7							
28	74.4	5.3	1152	12	CNS034DK							
29	74.2	5.3	1187	17	BC309087							
30	74	5.3	928	17	B11102							
31	73.2	5.3	1161	17	CNS00DKY							
32	73.2	5.2	923	17	CNS073Y8							
33	73	5.2	1225	17	AG058591							
34	73	5.2	1225	17	CNS0161D							
35	72.6	5.2	1101	17	CNS001FB							
36	72.4	5.2	734	17	CNS010MP							
37	72.4	5.2	961	17	CNS008HT							
38	72.2	5.2	994	17	CNS04NUT							
39	72.2	5.2	1006	17	CNS00813							
40	72	5.2	945	17	CNS04DOK							
41	71.8	5.2	976	17	CNS04ESM							
42	71.8	5.2	1101	17	CNS00EVL							
43	71.6	5.1	895	17	CNS00EVL							
44	71.4	5.1	867	17	CNS075BG							
45	71.4	5.1	923	17	AG058591							

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of Rpci-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL069706
GSS...
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named Rpci-98 and was constructed by partial

ECORI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1.1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_11b="RPC1-98"
/note="end : 17"

BASE COUNT 419 a 91 c 60 g 299 t 232 others

ORIGIN

Query Match 7.3%; Score 101.2; DB 17; Length 1101;
Best Local Similarity 35.7%; Pred. No. 3,5e-09;
Matches 223; Conservative 109; Mismatches 287; Indels 5; Gaps 1;

606 AAGTCGACGAATTAATCTGCTTAGAGCTATTTAGACGATTAATTAACATTAACCT 665

483 AAAATTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 542

666 AGATAGATGTAATTAATAATCAATGCTTCTATTGATTAATTTAGCATTTTGCAA 725

543 AAAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 602

726 AGCATTAATCCAAAGAGATGGAATAATTAACAAATTAATTAATTAATTAATTA 785

603 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 662

786 AAGAGACTTCTGGGAAACCCCTTACGATATGATACAGATTAATTAATTAATTAAT 845

663 WATTWMTAATTAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 722

846 TTCTAGTCTTAAGATGTTCAATGAAATAATTAACAAATTAATTAATTAATTAAT 905

723 AAWTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 782

906 GCCATCTATCTAAGGAAATGGAATTAATTAATTAATTAATTAATTAATTAATTA 965

783 WTATWMTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 842

966 AATTATTAATAAAGATTAATACACCTAATTAATTAATTAATTAATTAATTAATTA 1025

843 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 897

1026 TTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1085

898 WATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 957

1086 TGGAAATGCTTTAATTAATCTTGAATGAATTTCTAAGAGTATTAATCCCGAGTAT 1145

958 TATWMTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1017

1146 CCCCTCTTTAATAAATAAAGAGCAAGTAATTAATTAATTAATTAATTAATTAATTA 1205

1018 TATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1077

1206 ACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1229

1078 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1101

RESULT 2
CNS0039G/c 1101 bp DNA linear GSS 03-JUN-1999

LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #

BAC08K10 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL063921
VERSION
AL063921.1 GI:4941778

KEYWORDS

GSS.
Drosophila melanogaster.

SOURCE

Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

ORGANISM

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1101)
Genoscope.

AUTHORS

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;

TITLE

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

JOURNAL

- Web : www.genoscope.cns.fr

COMMENT

determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

FEATURES

The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila

source

1.1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_11b="BACR08K10"
/clone_11b="RPC1-98"
/note="end : TET3"

BASE COUNT

201 a 64 c 131 g 202 t 503 others

ORIGIN

Query Match 6.7%; Score 92.6; DB 17; Length 1101;
Best Local Similarity 20.6%; Pred. No. 1.4e-07;
Matches 139; Conservative 295; Mismatches 228; Indels 13; Gaps 3;

433 GGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 492

1101 GAKRWGDDTMDWRKDDDDDKKKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 1042

493 AGGATTTACCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 552

1041 WWWWMTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 982

553 ACTAATGATAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 612

981 AGRDGGRRRRKDDDDDDDKKKKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 923

613 GAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 672

922 KADDGACDDKDDGKXADDDTGGTDDKDDKDDKDDKDDKDDKDDKDDKDDKDD 863

673 TGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 732

862 GGVADADWTTWPAADWADDDWADWADWADWADWADWADWADWADWADWADWADW 803

733 AATCCAAAGAGAGTTGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTA 792

802 RDRRRADRRDAADDDDAATTTT-----TTTTRBDDDKWKTDTWTBAAADRTWD 750

793 TTCTGGGAAACCTTTACGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 852

749 RDDDDDRBRAGTAGRRWRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 690

853 TCTAAGAGTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 912

689 TGRRRRRRRRATWDRDADWADWADWADWADWADWADWADWADWADWADWADW 630

QY 913 TATACAGCAAAATGATATATATATAGAGCTTATTAATGACCTAAATTTATT 972
 Db 629 WTWMAWMAWMDWTRADRDWADTWTDARKADROWAKARAWBRDRBAABADRW 570
 QY 973 ATAAAAAGATATACACCTAATTAATGAATAGATCTTTGTAATCAGGTCATTTATT 1032
 Db 569 TTGKTTTATTT-----TWAKAAWMAWMAWATTTATTTTWTWTWTWTWTWTWT 515
 QY 1033 AAATTTATATATATATATTAACAATTAATGAGCACTTGTAGCTTATCCGAAGATGGAAT 1092
 Db 514 TAAWMAWMAWMTATWAAWTAAMAAAAAATTTTTTTTTTTTAAWMAWMAWMTWT 455
 QY 1093 GCCTTAAATATCTT 1107
 Db 454 TTWTTTTWMAATTTT 440

RESULT 3
 CDS020N7 1190 bp DNA linear GSS 14-MAY-2000
 LOCUS Tetraodon nigroviridis genome survey sequence PUC-01 end of clone
 DEFINITION 153P04 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL206908.1 GI:7865727
 VERSION GSS: genome survey sequence.
 KEYWORDS Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 1190)
 Bernot,Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished

TITLE 2 (bases 1 to 1190)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billaule,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished

JOURNAL 3 (bases 1 to 1190)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

REFERENCE 1 (bases 1 to 1190)
 Bernot,Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished

TITLE 2 (bases 1 to 1190)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billaule,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished

JOURNAL 3 (bases 1 to 1190)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

REFERENCE 1 (bases 1 to 1190)
 Bernot,Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished

TITLE 2 (bases 1 to 1190)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billaule,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished

JOURNAL 3 (bases 1 to 1190)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

REFERENCE 1 (bases 1 to 1190)
 Bernot,Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished

TITLE 2 (bases 1 to 1190)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billaule,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished

JOURNAL 3 (bases 1 to 1190)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

REFERENCE 1 (bases 1 to 1190)
 Bernot,Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished

TITLE 2 (bases 1 to 1190)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billaule,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Unpublished

JOURNAL 3 (bases 1 to 1190)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)

QY 684 TATCATACGTTCTTATGATTAATTTAGCATTTTGGCAAGCATTAAATCCAAAAGA 743
 Db 643 AAATTTAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 702
 QY 744 GATTGAAAATTTATACAGAGT-TATTTATCTATTAACCTTTTAAAGACTTCGGGAA 802
 Db 703 TAT 762
 QY 803 ACCCTTACGATATGATPACAGAAATATATTTAATACAGTAGCTTCTAGTTCTAAGATG 862
 Db 763 AATTTTNTWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 822
 QY 863 TTCATTTGAAAATATACAGATTTATGATTTGCAAAATGGCCATTCGATACG 922
 Db 823 TATTTTTTTTAT 882
 QY 923 GAAAATTTGAAAT 981
 Db 883 TAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 942
 QY 982 TATACACCTTAATTAATGAAATGATTTCTTTGTTAAATCAGGATTTTAAATATAT 1041
 Db 943 TAAATTTATWCAAT 1002
 QY 1042 GTATCATATACAT 1067
 Db 1003 CACAAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMAWMA 1028

RESULT 4
 CDS0167M 1201 bp DNA linear GSS 26-JUL-1999
 LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC
 DEFINITION BACN15M24 of Drosophila library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.
 ACCESSION AL106396.1 GI:5621701
 VERSION GSS.
 KEYWORDS Drosophila melanogaster.
 SOURCE Drosophila melanogaster.
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1201)
 Genoscope.
 Direct Submission
 Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage ;
 BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
 - web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a
 collaboration with the European Drosophila Genome Project (EDGP) -
 http://www.edgp.ebi.ac.uk - This Drosophila melanogaster BAC
 library (Dros BAC) was made by Alain Billard at CEPH (Centre
 d'Etude du Polymorphisme Humain) with funding provided by a MRC
 project grant. The DNA was prepared from embryos by Alain Bucheton
 and Genevieve Payan. It has been constructed in the vector
 pBelobAC11.

FEATURES
 source 1. 1201
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 /clone="BACN15M24"
 /clone_lib="DrosBAC"
 /plasmid="pBelobAC11"
 /note="end : T7"

BASE COUNT 323 a 87 c 79 g 551 t 161 others
 ORIGIN

Query Match 6.4%; Score 89; DB 17; Length 1201;
 Best Local Similarity 37.0%; Pred. No. 6,2e-07;
 Matches 273; Conservative 90; Mismatches 361; Indels 14; Gaps 1;

QY 636 TAGAGAGGATTAAT 695

Accession	Version	Keywords	Source	Organism	Reference Authors	Title
D1187	1					
Q1696	1					
D1127	1					
Q1756	1					
D1067	1					
Q1816	1					
D1007	1					
Q1876	1					
D1947	1					
Q1936	1					
D1887	1					
Q1996	1					
D1827	1					
Q1056	1					
D1767	1					
Q1116	1					
D1707	1					
Q1176	1					
D1661	1					
Q1236	1					
D1601	1					
Q1296	1					
D1541	1					
Q1356	1					
D1481	1					
RESULT 5						
CNS020N7/c						
LOCUS						
DEFINITION						
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						

JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1190)
AUTHORS	Roeck-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C., Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1190)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000)
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .
FEATURES	Location/Qualifiers
SOURCE	1. 1190
	/organism="Tetraodon nigroviridis"
	/db_xref="taxon:99883"
	/clone="153P04"
	/clone_1ib="G"
	/note="Genoscope sequence ID : COAG153DH02SP1-end :
	PUC-Ori"
BASE COUNT	373 a 229 c 141 g 312 t 135 others
ORIGIN	
Query Match	6.4%; Score 88.4; DB 17; Length 1190;
Best Local Similarity	39.7%; Pred. No. 8e-07;
Matches	Conservative 65; Mismatches 294; Indels 2; Gaps 2
Oy	GGGACAGCTTAGCAAAATACCTTTTAGGAGATTACCCTGATAATTAAGCTATATTAGC 527
Db	1040 GAKAGCGATRTATATNNWTKTGTRGTGWGKGKGTGAATATTATTGTTWKATATATA 981
Oy	528 AAATAAATGGTTTTTATTAACACTATTACTAATGATGATGATTAATTCCTTCGAATTTGGATAT 587
Db	980 TATTNATTTTATATATATAMTNATNTGTMTATATWTATATATATWTATATATATATKTAT 921
Oy	588 AAATGAGTACTTTAGGGAAGTGCGAAATTTACTGGTTTAGAGCATTAAGACAGATA 647
Db	920 ATATATATMTATATNTNMWNTATATANATTTTWTATATATTTTAAATAAATATTTT 861
Oy	648 TATATTAACATTTAAACTGATAGATGATATATATATTAATCAATACGCTTCTATGTGTA 707
Db	860 TATATATATATATKAATATAMANNATATAMAAAAAAMATAM-MWTATATTTATAAAMAANA 802
Oy	708 ATTTAGGATATTTTGCAGAACATTAATCCAAAAGAGATTGAAAATTTATACACAAGTTA 767
Db	801 WTAWTWAMTANTATATATATATATAMWTTWMANNAATATATATNAAAAAAAAAAAAATATA 742
Oy	768 TTATATCTATACCTTTTAAAGACATTCGCGGAAACCCCTTACGATGATGACAGATA 827
Db	741 ATATATTTTATATATATATATATNTATATATATATATATATATTTATTTATTTATATAT 682
Oy	828 TTAATTAATACGATGCTTCTAGTTCTAAAGATGTTCAATGGAANAATATATACAGATTA 887
Db	681 TTTTWTATTTTATMANWTATATATATTTTWTWTAAATTTTATAMTATATTTATATAAT 622
Oy	888 TAGTATTTTACAAATCCGCCATCGTACTAAGCGAAAATGGAATTTATATTTATAGAAG 947
Db	621 TTTTWTTTTWTMMWMAATMANATATATVMAAAANNTTMAAAMAAATTAANAATTT-WAAAT 563
Oy	948 GTTATATAAAGCAATAATTTATTAATAAAGATATACACCTAATATGAATATGATTC 1007
Db	562 TTAATAAAAAATTTTAAATATATWTNNAAATNTNTATTAATAAAATWTJTAATTTT 503
Oy	1008 TTTTGTTAAATCAGCATTTTATTAATATATATGATCATATATACAAATATGAGCACA 1066
Db	502 TTTTTTTTTTTTTTTATTTATATTTTAAATTTTAAATTTTAAAMAAMAVVVMM 444

[illegible]

RESULT	8
CNSOECOT	
LOCUS	1101 bp DNA linear GSS 04-JUN-1996
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC:
	BACR29P01 of RPc1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.
ACCESSION	AL069440
VERSION	AL069440.1 GI:4949583
KEYWORDS	GSS.
SOURCE	Drosophila melanogaster.
ORGANISM	Drosophila melanogaster. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Epiphytoidae; Drosophilidae; Drosophila. 1 (bases 1 to 1101) Genoscope. Direct Submission Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segrete@genoscope.cns.fr - Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library.html was prepared by Kazutoyo Osegaawa and Airon Mammoter in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPc1-98 and was constructed by partial EcoRI digestion of named RPc1-DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the library pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

FEATURES	location/Qualifiers
source	1. .1101
	/organism="Drosophila melanogaster
	/db_xref="taxon:7227"
	/clone="BACR29P01"

```

/clone_lib="RPIC1-98"
/notes="end : TET3"
BASE COUNT      366 a      66 c      104 g      351 t      214 others
ORIGIN
Query Match      6.0%  Score 83.6; DB 17; Length 1101;
Best Local Similarity 35.3%; Pred. No. 6.2e-06;
Matches 189, Conservative 85; Mismatches 261; Indels 0; Gaps 0

```

OY	5599	TTATGGAAGTCACAAGAAATTACTGGTTTGAAGCGCATTTAGACAGATATATATTAACCT	658
Db	494	TAATTAAMWTTTAAAAAAMWAAATAATAAATCAAMWTAATTAATTTWTTATTAATAA	553
OY	659	TAAACCTGATGATGTATATATATATCAATACGCTTTCTATGTGATAATTTAGGATAT	718
Db	554	MMWAKRTTTTTATWMAATATAAAAAAAWAAATTTATTTTATTTTATATTAATAAGWMAA	613
OY	719	TTTGGCAAGCATTAATATCCAAAAGAGATTGGAATAATTTACACAGATTATATCATATA	778
Db	614	AAAAAMWRTATAAMHWTAATAATTATATWTTMAMATTTTAAATMATTATTTAAMWTAAT	673
OY	779	CCTTTTTAAGAGACTTCTGGGGAAACCTTTACGATATGATACAGAAATTTATTAATAC	838
Db	674	WTTTAAAMWTTTAAWTATATAAAARWMAATTTWMAAAAAAAWTTTWAATTAAMAAAT	733
OY	839	CAGTACGCTTGAGTTCTAAGATGTCAATTGAATAAATAATACATATATATGTATTTGA	898
Db	734	AAATTTTTTTTATTTAAAMWATTAAAAATTANTWATWAAAMWTTWMAAAAAAATFNMTA	793
OY	899	CAATGCCGCATCGTATCTAACGGAATAATTGAATATATATTTATAGAAGCTTATATAATG	958
Db	794	TAAATATWMAKATTAATWATATAMWTTTAAATATWATRTAAATTTWTAATWTTTTTTTAAW	853
OY	959	GACATAAAATTTATATAAAGAATATACACCTAATAAGAATACATCTTTGGTAAAT	1018
Db	854	TTTAAATTTTTTTAAWTAAMWATGTAATWAAATTAAMWTTTAAAAAAMAATGWTTTTAAAT	913
OY	1019	CAGGTGATTTTATTAATATATATATGATCATATTAACAAATAGACACATTTGAGTATATC	1078
Db	914	TWMCCTTAATTTTTTATTWMAAATGTATATAAAAAAMWATTATGATATMMWAAATRTWWTG	973
OY	1079	CGAAGATNGGAAGCCCTTATATAATCTTGATAGAACTCTAAGAGTACGTTATATA	1133
Db	974	TATTTKAAATATWGAATATTTKAAAATAAAATATWTAAMAAARTKTGTRTWKWA	1028

RESULT 9	CNS0155H	1001 bp	DNA	Linear	GSS 26-JUL-1999
CNS0155H					
LOCUS					
DEFINITION	Drosophila melanogaster genome survey sequence S76 end of BAC BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL105023				
VERSION	AL105023.1	GI:5617037			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephyroidae; Drosophilidae; Drosophila. 1 (bases 1 to 1001) Genoscope.				
REFERENCE					
TITLE	Direct Submission				
AUTHORS	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :				
JOURNAL					

JOURNAL
Submitted (23-JUL-1999) genome - Centre National de Séquençage
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> . This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (centre
d'Etude du Polymorphisme Humain) with funding provided by a MC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector

[illegible]

please see <http://www.fruitfly.org> The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammeter in Pleier de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila/bac.htm>.

[illegible]

REFERENCE	TITLE	JOURNAL	REFERENCE	TITLE	JOURNAL
1 (bases 1 to 829)			2 (bases 1 to 829)		
Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Benoit,A., Fitzames,C., Wincker,P., Brothier,P., Quetier,F., Saurin,W. and Weissenbach,J.			Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Benoit,A. and Weissenbach,J.		
Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence			Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis		
Unpublished			Unpublished		
3 (bases 1 to 829)			3 (bases 1 to 829)		
Genoscope.			Genoscope.		
Direct Submission			Direct Submission		
Submitted (12-APR-2000)			Submitted (12-APR-2000)		
This sequence is a single read and was generated as part of a large scale clone and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .			This sequence is a single read and was generated as part of a large scale clone and sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .		

FEATURES					
source	Location/Qualifiers				
	1..829				
	/organism="Tetrarodon nigroviridis"				
	/db_xref="taxon:99883"				
	/clone="036f11"				
	/clone_id="G"				
	/note="Genoscope sequence ID : COHG036CC06SP1-end :				
	puc-ori"				
BASE COUNT	594 a	0 c	2 g	232 t	1 others
ORIGIN					
Query Match	5.7%;	Score 79.4;	DB 17;	Length 829;	
Best Local Similarity	46.1%;	Pred. No. 3.9e-05;			
Matches 337;	Conservative 0;	Mismatches 391;	Indels 3;	Gaps 2;	
Oy	499 TTACCTGATTAATTAAAGCTTATTACCACAATAATGGGTTTTTATTAACACTATTACTAAT	558			
Db	95 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	154			
Oy	559 GATAGATTATTCCTCGCAATTTGTATTAATGAAGTACTATGGGAAGTGACAAATT	618			
Db	155 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	214			
Oy	619 ACTGTTTAGAGCCTATTAGAGAGATTAATATATACATTAAAACCTAGATAGATGTAT	678			
Db	215 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	274			
Oy	679 AATATATCATATACGTTCTATTGATTAATTTAGGATATTTTGCAAGCATTAAATCCA	738			
Db	275 AATA--AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	332			
Oy	739 AAAGATTTGAAAAATTTATCACAGAATTATTATCTATACCTTTTAAAGACTTCG	798			
Db	333 TAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	391			
Oy	799 GGAAACCCCTTACGATATGATACAGATATTAATTTAATACAGTAGCTTCTAGTTCTAAA	858			
Db	392 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	451			
Oy	859 GATGTCATTAATGAAATAATTAACAGATTATATGTATTGACCAATGGCCCATCGPATCT	918			
Db	452 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	511			
Oy	919 AACGGAAATTGAATATATATATATAGAAAGTTATATAATGACCTAAAATTTATTTATAAA	978			
Db	512 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	571			
Oy	979 AGATATTACGCCAATTAATGAATAGAAATGACTCTTTGTTAATACGAGATTTTATTTAAATTA	1038			
Db	572 ATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	631			

[illegible]

RESULT 12	
AQ026918/c	
LOCUS	890 bp DNA linear GSS 30-JUN-1998
DEFINITION	CIT-HSP-2322B22.TF CIT-HSP Homo sapiens genomic clone 2322B22, DNA sequence.
ACCESSION	AQ026918
VERSION	AQ026918.1 GI:3267140
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 890)
AUTHORS	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL COMMENT	Unpublished (1998) Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0200 Fax: 301 838 0208 Email: mdamads@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tcdb/humgen/bac_end_search/bac_end_search.html . Seq primer: M13-21 Class: BAC ends.

[illegible]

us-09-816-467-3.rst

us-09-816-467-3.rst

us-09-816-467-3.rst

